International Journal for Housing Science and Its Applications

Publish September 9, 2025. Volume 47, Issue 1 Pages 279-291

https://doi.org/10.70517/ijhsa47123

Optimization of the model of project-based teaching of physical education courses in higher vocational colleges and universities based on the matrix arithmetic big ideology perspective

Yunfan Zhang[1,*] and Lingjuan Deng[1]

[1] School of Marxism, Henan Open University, Zhengzhou, Henan, 450003, China

Corresponding authors: (e-mail: zhjz119@163.com).

Abstract Based on the concepts of matrix arithmetic and big Civic and political education, this paper carries out optimization research on the project-based teaching mode of physical education courses in higher vocational colleges. At the level of teaching design, a teaching mode based on the programization of physical education courses is constructed, and the Civic and Political elements are integrated into the whole process of sports skills training. The inter-frame difference method is used to realize motion detection, and the Fourier transform technology and kernel component analysis are used for feature extraction and dimensionality reduction of the image. Then the hierarchical spatial feature extraction model CNN and temporal feature processing model LSTM are introduced to establish a recognition model based on CNN-LSTM hybrid neural network to recognize sports actions. The results show that the CNN-LSTM hybrid model is less accurate than the combined model in terms of performance metrics compared to the model when only CNN or LSTM alone module is used, and the error value of the combined model for action capture is lower, which proves that the combined model has stronger performance. In addition, the model in this paper is able to capture the spatio-temporal relationship between time series and joints more accurately and quickly, and in the process of realizing the evaluation of the standard degree of different actions, the action recognition error of the CNN-LSTM model is reduced by more than 20% compared with that of the traditional model, which shows that the model in this paper has an excellent recognition effect on different sports actions as well.

Index Terms Inter-frame differencing, Fourier transform, kernel component analysis, CNN-LSTM, sports action recognition

I. Introduction

With the curriculum reform and the upgrading of students' needs, carrying out high-level physical education program teaching can not only well supplement the defects of the existing teaching mode in higher vocational schools, but also effectively promote the comprehensive growth of students' physical and psychological qualities, which is of positive significance for the deepening of the school's high-quality parenting work [1]-[3]. In view of this, higher vocational physical education teachers should widely explore the scientific sports teaching mode, integrate the project-based thinking and methods into daily teaching practice, enhance students' sports awareness, and realize the multi-dimensional development of sports teaching. The so-called project-based teaching refers to a relatively independent project under the guidance of the teacher, which is completely handed over to the students to deal with independently, and the students, by carrying out the project, understand and grasp the whole process and the basic requirements in each link [4]. Accordingly, the key to project-based teaching lies in the rational design of projects and the cultivation of students' ability to handle projects. Its most significant feature is "project-based, teacher-guided, student-oriented", to realize the effective integration of project elements, so as to enhance the relevance of "teaching and learning", and to improve the effectiveness of teaching and learning and the ability of independence [5]-[7]. Project teaching method is conducive to teachers to better break down the content of physical education, and effectively guide students to find their favorite sports, driven by interest to explore the curriculum, to achieve the fine control of teaching and learning, and gradually improve and perfect their professional ability, critical thinking ability, practical ability and comprehensive ability [8]-[11].

In addition, in physical education teaching, teachers can create physical education classroom teaching projects, make it organic with competition activities, let students participate in the project competition and competition, master sports skills in the confrontation, improve the quality of sports, and promote the diversification of the form of physical education teaching [12]-[14]. At the same time, in 2022, the state clearly pointed out that the ideological and political education should be integrated into the cultivation of various types of skilled personnel, based on the ideological and political perspective can be seen in the new era of educational development requirements is that physical education and moral education must be further integrated, so that physical education teaching work can convey the idea of moral education [15]. And "big ideology and politics" is an important cornerstone of the development of higher vocational education, it has a very high applicability and significance in the construction of physical education system, providing a unique perspective for reform and innovation [16]. However, in the actual teaching mode of physical education program, the teaching design of the integration of teaching objectives and ideological education is only presented in the form of splicing, and there is not much correlation between them, which leads to students still emphasizing physical education over ideology. And the course evaluation process is even more experience-based, the effectiveness of project-based teaching did not have an in-depth understanding and improvement, with a fixed template, set for all types of sports [17]-[19]. Based on this, in order to better achieve the higher vocational sports courses in line with national policy, the need to optimize the sports project-based teaching mode.

In this paper, the teaching process is divided into five steps, namely, "creating a project - making a plan - exploring activities - communicating results - evaluating and summarizing", arranging the teaching content from the perspective of students, and starting from the three aspects of project preparation, implementation and summarization. Preparation, implementation and summary of the three aspects, to build a project-based teaching model based on the physical education curriculum in higher vocational colleges and universities. The inter-frame difference method is used to detect the students' sports movements, the Fourier transform technique is used to extract the sports movement features, and then the kernel component analysis method is used to select the important sports features. Matrix operation is then used for feature processing, and CNN model with hierarchical spatial feature extraction capability and LSTM method with temporal feature processing capability are introduced to construct an LSTM-CNN model that retains detailed features and extracts abstract features, so as to enhance its accuracy and robustness for action recognition. Finally, the recognition performance and application effect of the model are analyzed by examples.

II. The model design of project-based teaching of physical education courses in higher vocational colleges and universities

II. A. Design of Project-based Teaching Mode for Physical Education Courses in Higher Education Institutions

II. A. 1) Project-based teaching theory model design

Different teaching modes reflect different teaching purposes. In the process of teaching, teachers will adjust the teaching mode and teaching links according to the teaching needs. In this paper, according to previous reports, the teaching process is divided into five steps, namely, "creating a project - making a plan - exploring activities - exchanging results --Evaluation and summarization". The overall model of project-based teaching in physical education courses in higher vocational colleges is shown in Figure 1.

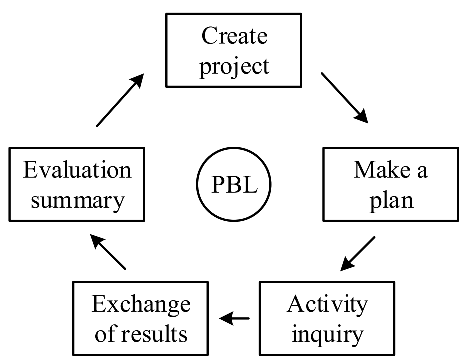


Figure 1: The overall model of the teaching of physical education

II. A. 2) Selection of teaching content

The textbook should be from the perspective of students, close to their lives, and the cases chosen can inspire students in higher vocational colleges to explore the physical education curriculum programs in their lives. In addition to theoretical knowledge, the textbook also has a lot of practical parts, which can make students understand the sports teaching mode of higher vocational colleges and universities more deeply through hands-on practice. While improving students' knowledge of physical education programs, it can stimulate students' interest in learning physical education programs, and then increase their enjoyment of physical education.

II. A. 3) Project-based teaching process design for physical education courses

In the teaching process, we divided the five steps in the design of the teaching theory model into three stages, namely, project preparation, project implementation and project summary. In the project preparation stage, there are creating projects and making plans, in the project implementation stage, there are activity exploration and results exchange, and in the project summary, there are evaluation and summarization. The specific contents are as follows:

(1) Project preparation stage

The project preparation stage is the most important first step in project-based teaching. It can stimulate students' participation in the classroom and the enthusiasm of the research project, and create a good start for the whole project-based teaching.

Teachers' activities: Teachers need to group students into small groups to ensure that each child has a group, and to clarify the division of labor within the group; teachers need to consider the degree of difficulty of the teaching content according to the students' situation; teachers need to create situations according to the project and the students' situation.

Student activities: Students are the main body of project-based teaching. First of all, students need to make a clear division of labor, understand their own tasks and responsibilities in the group, take up their own responsibilities in the group, and provide their own share for the subsequent group to complete the project.

(2) Project implementation stage

The project implementation stage is the main body of the whole project-based teaching, in this process, the students independently study and explore, with the help of various learning resources to obtain knowledge and skills enhancement.

Teacher activities: Teachers are the supervisors and organizers of learning in the project implementation stage. Teachers should make clear the project, set up the task; make a good predetermination of the students' problems, classify the problems as well as prepare teaching materials in advance, and analyze and explain these core problems.

Student activities: students in the project implementation stage according to the pre-planning, the group to carry out the implementation of sports projects, collaborative cooperation, and common completion of the project sports project. In the display and exchange of sports projects, students need to demonstrate their sports projects, introduce the knowledge contained in the sports projects, sports projects with the highlights as well as shortcomings and directions for improvement.

(3) Summarization stage

In the project summarization stage is mainly the process of evaluation. Compared with the traditional single teaching evaluation, the optimized project-based teaching mode of higher vocational colleges and universities has various forms of evaluation, such as "outcome evaluation, process evaluation, teacher evaluation, mutual evaluation and self-assessment".

Teachers' activities: Teachers need to evaluate and summarize the students' sports projects, and also need to give process evaluation to the students' performance in the whole project-based teaching, and point out the solutions and ways to solve the problems encountered by the groups in the process of activities, so as to provide good guidance for the subsequent project-based teaching of sports courses.

Student activities: intra-group communication and inter-group communication can sublimate the project, each group according to the performance of the evaluation, you can point out the problems of other groups, you can also learn from the advantages of other groups, for the next project to do better.

II. B. Recognition model for sports movements

II. B. 1) Sports movement detection

In the process of sports action recognition, the first step is to detect the student's action, combined with the characteristics of the student's action, the inter-frame difference method [20] is used to realize the action detection, and the detection results are processed such as expansion, corrosion contour enhancement, etc., as follows.

(1) Let $f(i,j)$ be the frame image of the sports video, and the two frames before and after are $f_i(i,j)$ and $f_{i+1}(i,j)$, and $Gray(f_1(i,j))$ and $Gray(f_2(i,j))$ are used as the grayscale images, and the transformation is specified as follows:

$$Gray(f(i,j)) = 0.11R(i,j) + 0.59G(i,j) + 0.3B(i,j) \tag{1}$$

(2) A threshold ε is used to judge the noise in the sports image and the following equation is used to perform a difference operation to obtain a binary image $D(i,j)$, if $D(i,j)=1$ it means that the pixel position is the position where the action is generated. I.e:

$$D(i,j) = \begin{cases} 0, & |f_i(i,j) - f_{i+1}(i,j)| < \varepsilon \\ 1, & |f_i(i,j) - f_{i+1}(i,j)| \geq \varepsilon \end{cases} \tag{2}$$

(3) The expansion and corrosion process of the binary image is:

$$D = X \oplus S = \{(x,y) \mid (x,y) \in X, S(x,y) \cap X \neq \varnothing\} \tag{3}$$

$$E = X \otimes S = \{(x,y) \mid (x,y) \cap X, S(x,y) \subseteq X\} \tag{4}$$

(4) Gaussian filtering is used to remove the noise from the binary image and the Gaussian function is defined as follows:

$$f(x,y)=\frac{1}{2\pi\sigma^2}e^{(-\frac{x^2+y^2}{2\sigma^2})} \tag{5}$$

II. B. 2) Extracting sports movement characteristics

The contour line is obtained by taking the edge point of the student's head as the starting point of the contour and dividing it into M segments of arcs to obtain the contour of one's running action. Assuming that the sports action contour is $R=(r_1,r_2,\cdots,r_N)$, and using Fourier transform for processing, we have:

$$a_n=\frac{1}{N}\sum_{i=1}^{N}r_i\exp(-j2\pi ni/N) \tag{6}$$

Normalizing the Fourier coefficients yields a Fourier descriptor of the sports action feature as:

$$f=\left[\frac{\|a_2\|}{\|a_1\|},\frac{\|a_3\|}{\|a_1\|},\frac{\|a_4\|}{\|a_1\|},\cdots,\frac{\|a_N\|}{\|a_1\|}\right] \tag{7}$$

When 64 low-frequency feature components are selected as sports action recognition features toward the manifest, there are:

$$\begin{aligned}F&=\left[\frac{\|a_2\|}{\|a_1\|},\frac{\|a_3\|}{\|a_1\|},\frac{\|a_4\|}{\|a_1\|},\cdots,\frac{\|a_{64}\|}{\|a_1\|}\right]\\&=[f_1,f_2,f_3,\cdots,f_{64}]\end{aligned} \tag{8}$$

II. B. 3) Selection of important features for nuclear component analysis

If 50 sports action features are directly input to the classifier for recognition, it is computationally intensive and feature redundancy, which interferes with sports action recognition, for this reason KPCA [21] is used to select the important features.KPCA is a nonlinear feature selection method, which performs a nonlinear mapping of the original features by means of an introductory kernel function, and then extracts the main features. Let the training set be $X=\{x_1,x_2,\cdots,x_n)$ and ϕ denotes the nonlinear mapping, then the covariance matrix is given by:

$$C=\frac{1}{M}\sum_{j=1}^{N}\phi(x_j)\phi^{\text{û}}(x_j) \tag{9}$$

Decomposition of c yields: $\lambda v=Cv$, v consists of $\phi(x_1),\phi(x_2),\cdots,\phi(x_n)$ and there is:

$$\lambda(\phi(x_i),v')=\langle\phi(x_i),Cv'\rangle,k=1,2,\cdots,M \tag{10}$$

Since v' is a linear mixture of $\phi(x)$, one obtains:

$$v'=\sum_{i=1}^{N}c_i'\phi(x_i) \tag{11}$$

Let $k_{ij}=\langle\phi(x_i),\phi(x_j)\rangle$, which follows from the above equation:

$$M\lambda'c'=Kc' \tag{12}$$

The eigenvectors with eigenvalues greater than 0 are $c^p,c^{p+1},\ldots,c^n$, and the projection of the sample $\phi(x)$ on c' is:

$$\phi(x_i)=\phi(x_i)-\frac{1}{M}\sum_{i=1}^{N}\phi(x_i) \tag{13}$$

All projections form a vector $g(x)=[g_1(x),g_2(x),\cdots,g_i(x)]^T$, i.e., the main features are extracted. Using $K_1(x_i,x)=\langle\phi(x_i),\phi(x)\rangle$, the projection of the k th principal element v^t of x is:

$$g(x)=\langle v',\phi(x)\rangle=\sum_{i=1}^{N}c'K_1(x_i,x) \tag{14}$$

II. C. Image processing and matrix operation after sports action recognition

Image processing involves a large number of matrix operations, which can be understood through image transformations. The following discusses the relationship between transposition of chunked matrices, matrix addition and subtraction operations, number multiplication, inverse of matrices, etc. and image transformations.

Let the chunking matrix $M=\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$, then $M^T=\begin{pmatrix} M_{11}^T & M_{21}^T \\ M_{12}^T & M_{22}^T \end{pmatrix}$.

Number multiplication of matrices - changing the brightness of an image. Multiplying a matrix by a positive number shows up on the image as a change in overall brightness. When this number is less than 1, it is shown as a decrease in overall brightness; when this number is greater than 1, it is shown as an increase in overall brightness; and when it is equal to 1, it is shown as no change in brightness.

Matrix subtraction and number multiplication operations - change the contrast of the image. When the difference between the values in the image is small, its contrast is small, which can be adjusted by matrix subtraction and multiplication. A simple formula for adjusting the contrast of an image is:

$$M_{m\times n}^{'}=\frac{1}{h_{\max}-h_{\min}}\times(M_{m\times n}-h_{\min}\times I_{m\times n}) \tag{15}$$

where $m\times n$ denotes the size of the image matrix, $h_{\min},h_{\max}$ denote the values of the smallest and largest elements of the matrix, respectively, and $I_{m\times n}$ denotes the matrix with the same size as the image and with all elements 1.

Subtraction of matrices - the negative of an image. The photographic film is so different from the actual image that what is bright in actuality appears dull on the film, and what is dull in actuality is bright on the film, that it is called a negative or negative. To wit:

$$M_{m\times n}^{-}=\alpha\times I_{m\times n}-M_{m\times n} \tag{16}$$

α is the product coefficient, $\alpha=255$ when the elements in $M_{m\times n}$ take values in the range [0, 255], and $\alpha=1$ when the elements in $M_{m\times n}$ take values in the range [0, 1].

Matrix addition and number multiplication operations - adding a watermark to an image. To wit:

$$M_{m\times n}^{'}=M_{m\times n}+\beta\times N_{m\times n}^{-} \tag{17}$$

where $M_{m\times n}^{*}$ denotes the watermarked image matrix, $N_{m\times n}^{-}$ denotes the watermarking matrix, and $\beta\in(0,1]$ denotes the coefficient of the watermark intensity. Take β =1, 0.75, 0.5, 0.25 respectively.

Role of Invertible Matrices - Simple Image Encryption and Recovery. Invertible matrices have an important role in linear algebra and invertible matrices play an important role in image processing. By multiplying an invertible matrix on the left-hand side of the corresponding matrix of an image and compressing the resulting matrix into a displayable interval, one has:

$$Y_{m\times n}=A_{m\times m}M_{m\times n},Y_{m\times n}^{'}=\frac{\alpha}{h_{\max}}Y_{m\times n} \tag{18}$$

where $Y_{m\times n}$ denotes the result of the product of the invertible matrix $A_{m\times m}$ and $M_{m\times n}$, $Y_{m\times n}$ denotes the compression of the elements in the image into the interval $[0,\alpha]$, $h_{\max}$ denotes the largest element in $Y_{m\times n}$, and α has the same meaning as before.

If this process is reversed, i.e:

$$M_{m\times n}=A_{m\times m}^{-1}\frac{h_{\max}}{\alpha}Y_{m\times n} \tag{19}$$

Here, $A_{m\times m}^{-1}$ denotes the inverse matrix of the invertible matrix $A_{m\times m}$. Even if $h_{\max}$ is not known, it is possible to get a restoration result that is basically the same as the original image by adjusting the contrast.

II. D. Methods for recognizing actions in sports training videos

II. D. 1) Spatial feature extraction of actions based on CNN layer

The first CNN layer extracts the low-level spatial features of the actions in the sports training video in the following steps.

Step 1: Assuming that the sample of the sports training video is x, the convolutional layer uses the convolutional kernel $\tilde{\omega}$ to perform convolutional operations on x, and automatically learns and extracts the local low-level spatial features of x. The formula for the first CNN layer is as follows:

$$z(i,j)=\sum_{h=1}^{H}\sum_{w=1}^{W}x_{i+h,j+w}\times\bar{\varpi}_{h,w}+b \tag{20}$$

Wherein, the local low-level spatial feature map z of the action in the sports training video is located at (i,j), the height and width of $\tilde{\omega}$ are H and W; $h\in H$, $w\in W$, b is the bias, and the value of the $i+h$ th row of the $j+w$ th column within x is $x_{i+h,j+w}$.

Step 2: The pooling layer performs a downsampling operation on $z(i,j)$ with the aim of reducing the size and parametric number of the local low-level spatial feature maps of the actions in the sports training video, improving the training speed and preventing overfitting. In the processing of the sports training video, the l th local low-level spatial feature map of the action output by this layer is:

$$z_l^{'}=\varphi(\alpha_l down(z_l(i,j))+b_l) \tag{21}$$

where φ is the modified linear unit (ReLU) activation function; $do\varpi n(\cdot)$ is the maximum pooling function, α_l is the pooling weights, and α_l is the pooling bias.

Step 3: The fully connected layer integrates the local low-level spatial features of the actions extracted in Step 2 to form global low-level spatial features. The integration formula for the fully connected layer is as follows:

$$\hat{z}^n=\varphi(\alpha^n z_i^{'}+b^n) \tag{22}$$

where is α^n and b^n are the weights and biases of the fully connected layer.

Step 4: The output layer outputs the global low-level spatial features $\hat{z}$ of the actions in the sports training video obtained from the integration in Step 3.

Finally, the second CNN layer extracts the global high-level spatial features $\dot{x}$ of the actions in the sports training video within $\dot{z}$, with the aim of minimizing the loss of information and improving the richness of the global high-level spatial features.

II. D. 2) Timing feature extraction for LSTM layer based actions

In extracting the temporal features of actions in sports training videos,the LSTM has a comprehensive capturing ability. Let the input gate of the LSTM layer be s, s determines the portion of the input $\hat{x}$ at t time step that needs to be accepted by the g of the cell state. Inputting the global high-level spatial features $\dot{x}$ of the action in the sports training video within s yields, the

$$s_t=f(\omega_s(\hat{x}_t,y_{t-1})+b_s) \tag{23}$$

where ω, and b, are the weights and biases of s, f is the Sigmoid function, y_{i-1} is the hidden state, and y_{i-1} contains all the important information about the LSTM from the beginning of processing $\dot{x}$ to the time step of $t-1$, which is to say, it contains the action of the sports training video temporal features.

The forgetting gate of the LSTM layer is h, and h determines the global high-level spatial feature information of the actions in the sports training video that needs to be forgotten and needs to be retained in the $t-1$ time step of g, which is calculated as follows:

$$h_t=f(\omega_h(\hat{x}_t,y_{t-1})+b_h) \tag{24}$$

where ω_h and b_h are the weights and biases of h.

The output gate of the ISTM layer is o, and o determines the global high-level spatial feature information of the action in the sports training video that needs to be output to the current hidden state in g, which is calculated as follows:

$$o_t=f(\omega_o(\hat{x}_i,y_{t-1})+b_o) \tag{25}$$

where ω_o and b_o are the weights and biases of o.

The g is responsible for helping the LSTM layer to capture the long-term dependencies in $\hat{x}$, which is calculated as follows:

$$g_t=h_t\square\ g_{t-1}+s_t\square\ \tanh(\omega_g(\hat{x}_t,y_{t-1})+b_g) \tag{26}$$

where ω_g and b_g are the weights and biases of g.

The temporal features of the actions in the sports training video extracted by the ISTM layer are.

$$y_t=o_t^{\circ}\tanh(g_t) \tag{27}$$

The attention mechanism can help the LSTM layer pay more attention to the important time steps, improve the accuracy of temporal feature extraction of actions in sports training videos, get more valuable temporal features, and improve the recognition effect of actions in sports training videos.

This attention mechanism contains two parts, namely channel attention λ^p and spatial attention λ^q. Processing y_i using this attention mechanism yields:

$$y_i^p = \lambda^p(y_i) \otimes y_i \tag{28}$$

$$y_i^q = \lambda^q(y_i) \otimes y_i \tag{29}$$

where y^p and y^q are the temporal features of actions in sports training videos with channel attention and spatial attention, respectively.

The fusion of y_i^p and y_i^q can obtain more valuable temporal features of actions in sports training videos with the following formula:

$$\hat{y}_i = f(\varphi(AvgPool(y_i^p); MaxPool(y_i^q))) \tag{30}$$

where $AvgPool(y_i^p)$ and $MaxPool(y_i^q)$ are average and maximum pooling.

II. D. 3) Recognition of actions in sports training videos

The combination of LSTM and CNN effectively integrates the spatial and temporal processing capabilities of the two, CNN can extract the spatial features of the video frames, and LSTM can analyze the temporal sequence of these spatial features. This complementarity enables the LSTM-CNN model to comprehensively capture the spatial and temporal dimensional information in sports training videos, significantly improving the accuracy of action recognition.

The recognition model of actions in sports training videos can be established by using LSTM-CNN [22], [23]. The specific steps for action recognition in sports training videos are as follows.

Step 1: Within the input layer, the original sports training video samples are input.

Step 2: The 1st CNN layer extracts the low-level spatial features $\hat{z}$ of the actions in the sports training video.

Step 3: Combining the low-level spatial feature master $\hat{z}$, the 2nd CNN layer extracts the high-level spatial feature $\dot{x}$ of the action in the sports training video.

Step 4: Combining the high-level spatial feature $\dot{x}$, the ISTM layer extracts the temporal feature y_i of the action in the sports training video.

Step 5: Using the attention mechanism, the extracted temporal features y_i are filtered to obtain more valuable temporal features $\hat{y}_i$.

Step 6: Using Softmax classifier, the output layer processes the temporal features $\hat{y}_i$ to obtain the probability value that the input sports training video sample belongs to a certain action category with the following formula:

$$\rho(r_\tau) = \frac{\exp(r(\hat{y}_t))}{\sum_{\tau=1}^{M} r_\tau} \tag{31}$$

where r_τ is the action category in the τ sports training video, M is the number of action categories in the sports training video, and the action category corresponding to the maximum probability is regarded as the final action recognition result in the sports training video.

III. Sports action recognition model application effect and simulation case analysis

III. A. Performance validation of visual image-based action recognition system

III. A. 1) Performance test of action recognition system under visual image

In order to explore the comprehensive performance of the sports action recognition model proposed in this paper, the performance of the sports action recognition system under the LSTM-CNN model is uniformly tested by the same operation method during the experimental process in order to avoid the influence of the different hardware equipment and operation methods on the experimental results. In this paper, we choose LFDD public dataset, Human4.5M public dataset and traditional sports action (TPA) dataset. The experimental results of watermark intensity coefficient of the model are shown in Fig. 2. In order to compare the recognition accuracy of the test set under different watermark strength coefficients, the recognition accuracy varies significantly when the watermark strength coefficients are different. When the ratio is maximum 1.00, the overall recognition accuracy of the model is the lowest, and the growth is slow, when the number of iterations reaches 50, the model recognition accuracy only reaches 0.7938; when the watermark strength coefficient is set too small for 0.25, although the model recognition accuracy improves rapidly, when the number of iterations is more than 30 there is instability in the recognition accuracy, and ultimately there is a downward trend; when the watermark strength coefficient is set at 0.75, the overall recognition accuracy is excellent, there are fluctuations in accuracy in the pre-iteration period, but when the number of iterations exceeds 43, it tends to stabilize, and the final recognition accuracy reaches the optimum (95.96%). Therefore, the subsequent system performance test set the watermark intensity factor of 0.75.

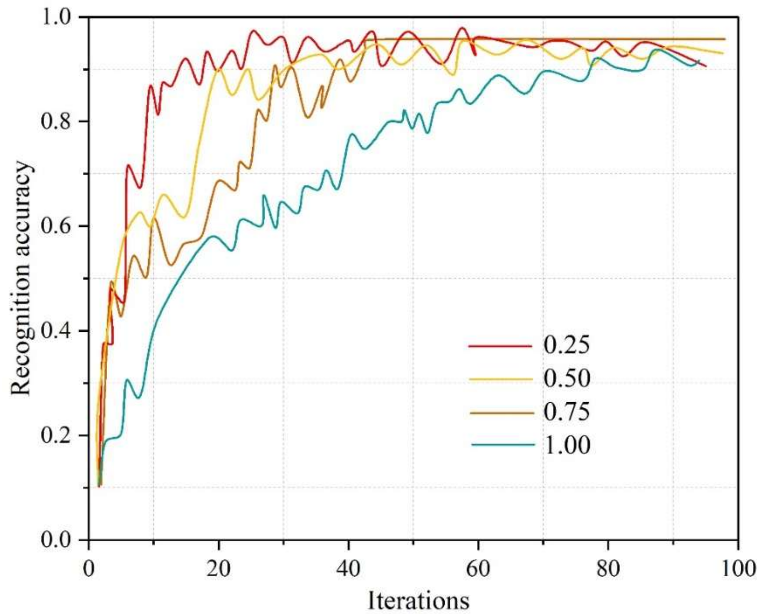


Figure 2: The results of the study model's β ratio

The loss curve of the model as a function of the TPA dataset is shown in Fig. 3. During the experiment, the research model was trained using the TPA dataset, and it can be seen that the training loss value gradually stabilizes and eventually reaches about 0.0147 when the number of iterations reaches 20 or more. The test loss value decreases rapidly in the first 16 iterations, converges gradually at around 36 iterations, and finally reaches a steady state at around 0.1251.

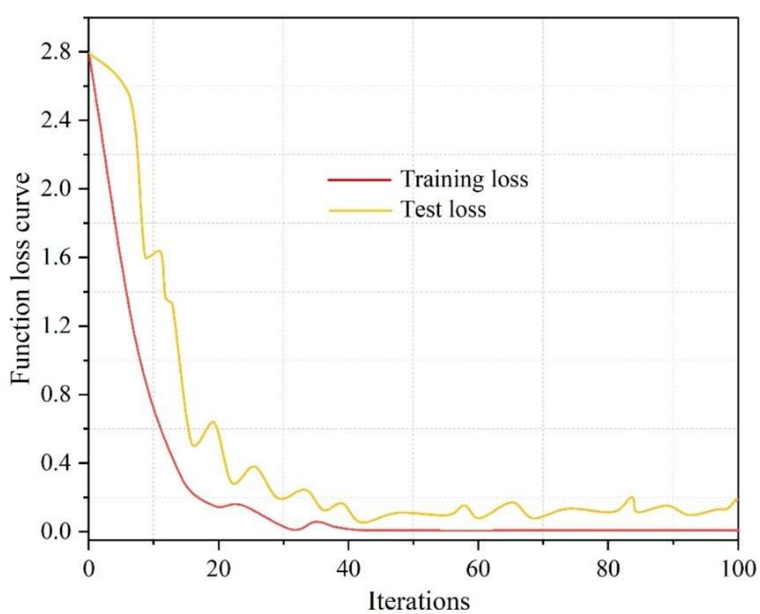


Figure 3: The function loss curve in the TPA data set

Figure 4 shows the test recognition accuracy and training recognition accuracy curves of the model in the TPA set. The test recognition accuracy tends to stabilize after 40 iterations, and the final recognition accuracy reaches about 0.9799; the recognition accuracy of the training set tends to stabilize after the 20th iteration, and maintains a steady state after 30 iterations, and finally converges to 0.9993.

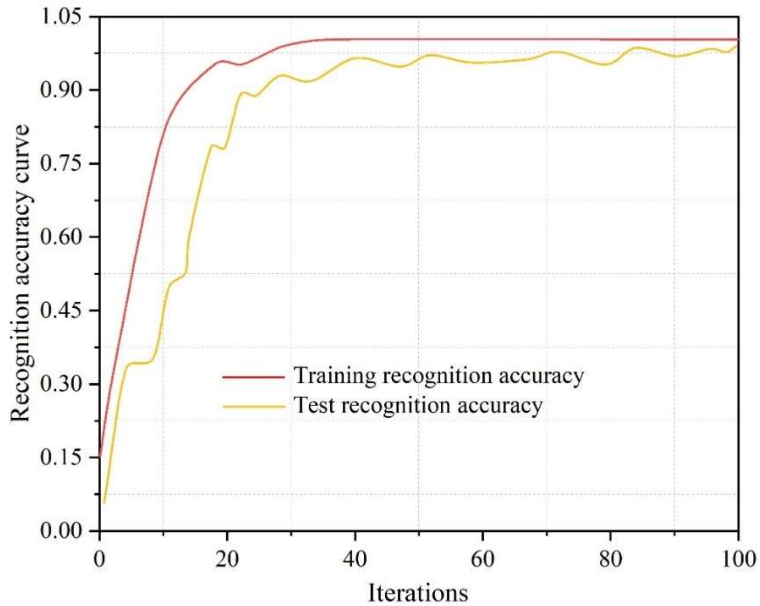


Figure 4: Test and training recognition accuracy curve of the model on TPA

III. A. 2) Ablation experiments

In order to verify the effect of LSTM module and CNN on the performance of the network model, a series of ablation experiments are first conducted. The precision rate (Pre), recall rate (Rec), F1 score (F1_score) and computational speed (CS) are used as the indexes for evaluating the performance strengths and weaknesses in the test model during the experiments, and the average value obtained after repeating the experiments for 10 times is used for each evaluation index.

The results of the model ablation experiments are shown in Table 1, which firstly shows that the models with the introduction of LSTM have different degrees of improvement in precision, recall and F1 score, while the computational speed is slowed down due to the introduction of the additional network structure; secondly, since the use of CNN can enhance the size of the field of view and reduce the loss of information, it has improved the performance of all the models to some extent; and lastly, compared to adding only the LSTM or CNN, the performance of the LSTM-CNN model combining the two is significantly improved, and its precision, recall, and F1 score are all greater than 99.9%, and its computational speed is increased by less than 1% compared to the two-layer CNN with the highest accuracy, which proves that the model in this paper is extremely effective in application.

Table 1: Model ablation experiment results

Insertion position	Pre(%)	Rec(%)	F1_score(%)	CS(s/epoch)
Conv-1	97.0352	97.0072	97.0147	1.1007
Conv-2	98.0262	98.9954	98.0219	1.0028
Conv-3	97.9994	97.9979	97.0386	1.0972
LSTM	97.9708	98.0067	97.9763	1.0999
CNN	98.1024	97.9892	97.9897	1.104
LSTM-CNN	99.9994	99.9012	99.9801	1.0091

III. B. Analysis of the model's effect on sports action recognition

The experiment was conducted on 50 students in the physical education classes of XX higher vocational colleges and universities. The experiment was set up with seven movements (hereinafter referred to as B1-B7), namely, left and right arm swing, up and down arm swing, push and pull, chest expansion, extension, jumping, and high-five), and for each movement, each test subject was required to repeat the performance three times to obtain 150 samples of movement data. 80% and 20% of the experimental samples were used as training samples and test samples, respectively. The experimental samples are preprocessed, normalized and prepared for use, and the neural network-based motion capture technology and inertial sensorbased motion capture technology are introduced for the same-condition comparison test to highlight the advantages of this paper's method in motion capture recognition.

III. B. 1) Motion Capture Accuracy Analysis

Students demonstrated sports movements, and three methods were used to capture the movement data of nine body parts of the test subjects: head, neck, left shoulder, right shoulder, waist, left wrist, right wrist, left foot, right foot, and the nine parts are abbreviated as A1-A9 in the following.Considering the data of the human body's joint points as the standard data, the sports movement capture methods based on CNN and LSTM were statistically analyzed to compare the accuracy of CNN and LSTM with the method of the present paper, LSTM- CNN is compared to analyze the motion capture accuracy of different methods. The statistical results of sports skill motion capture error are shown in Fig. 5. It can be seen that the method LSTM-CNN in this paper has the smallest error in the measurement of human movement, and the measurement error interval of 9 parts is [0.1584mm,0.4128mm], the error data fluctuates less, and the measurement results are more stable. Comparatively speaking, the error of CNN-based motion capture technology continues to rise, and the highest error can reach 2.6874mm, and the error fluctuation interval of LSTM-based motion capture technology is [1.169mm,2.2914mm], although the error fluctuation amplitude is not large, the measurement error value of this method is much higher than that of this paper's method, and the effect of the motion error capture is not ideal. It can be seen that the method in this paper shows the best effect of human movement measurement.

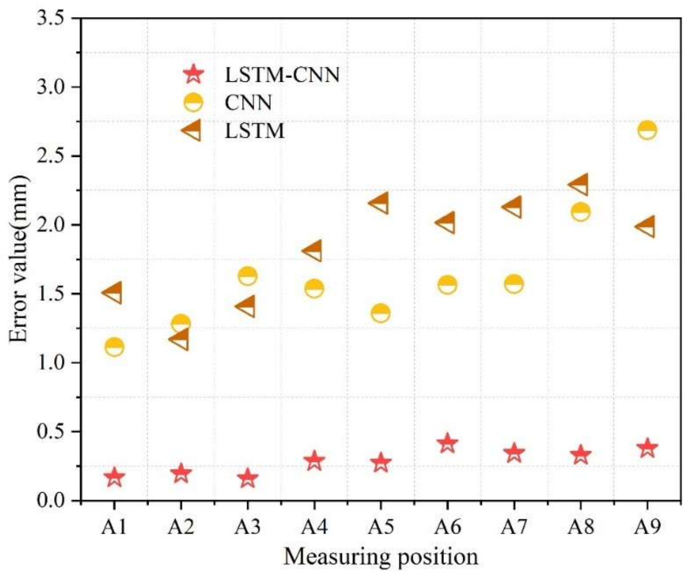


Figure 5: Statistical results of the motion capture error of sports skills

III. B. 2) Motion Recognition Accuracy Analysis

To further test the effectiveness of the three methods in recognizing the sports movements performed by the human body, the three methods were used to recognize the 30 experimental samples of movement data obtained. The statistical results of the recognition rate of sports movements captured by each method are shown in Table 2. It can be seen that the recognition rate of sports movements with the methods in this paper is the highest, and all of them are concentrated in more than 97%, and the highest recognition rate can reach 99.08%. The accuracy of the sports action capture model based on CNN method and LSTM method basically fluctuates between 80% and 90%, compared with this paper's method, or the accuracy rate is higher.

The maximum action recognition rate of CNN network-based motion capture technology is 90.07%, but only a few unfolding under the recognition of sports action accuracy is more satisfactory, the lowest recognition rate is only 86.89%. The LSTM-based motion capture technique shows the same pattern, with a minimum recognition rate of 83.05%. As the former applies the traditional neural network for action capture recognition, its weight determination is more random, which affects the efficiency and accuracy of action feature recognition, resulting in a low recognition rate of sports action; the latter focuses on the research of action capture time sequence feature processing, and applies the traditional support vector machine algorithm in behavior type recognition, although it can recognize the behavior type, but it results in poor stability and recognition accuracy is not ideal. Although the behavior types can be identified, the stability is poor and the accuracy is not satisfactory. The method in this paper combines the advantages of CNN and LSTM methods and is based on two-layer network, so that the human body's movements and behavior types can be identified with high accuracy.

Table 2: The results of the recognition rate of the physical action of the parties

Method	Recognition rate(%)						
	B1	B2	B3	B4	B5	B6	B7
LSTM-CNN	97.91	97.97	97.96	97.9	99.08	98.29	97.96
CNN	86.89	90.21	90.4	89.76	90.07	89.9	89.99
LSTM	83.05	88.11	88.06	88.27	88.27	88.41	87.58

III. C. Empirical analysis of visual image-based sports action recognition

III. C. 1) Comparison of Recognition Errors of Different Models for Sports Movements

For sports action recognition, it can enable teachers and students to understand part of the action key force point and movement mode from data and multidimensional level in a more comprehensive and three-dimensional perspective, and at the same time, to verify that the proposed action recognition and acquisition system can correctly follow the human body model for accurate capture of actions. In this study, CNN, CNN and LSTM are used as control models to test the results of multiple action recognition with the CNN-LSTM model proposed in this paper. The experimental subjects in this paper are 150 students in MG higher vocational colleges in the physical education classroom taijiquan teaching exams to recognize and analyze the sports actions in their classroom content presentations.

The results of the action recognition comparison test of different models are shown in Figure 6. It can be seen that the error accuracy of this paper's model can be kept below 33.5mm in the Leapfrog action and Squat action with serious occlusion, while the error accuracy of the GCN model reaches up to 45.48mm.Among the three comparative models, the best-performing model is the CNN-LSTM model, with an average error accuracy of 34.18mm, and the smallest error accuracy for the action is Sit-up, and the action Purch with the highest error precision reaches 37.02 mm; while the average error precision of LSTM model reaches 45.19 mm, and the average precision error of GCN model is more than 47.18 mm.The recognition accuracy of this paper's model is improved by more than 20% compared with the three comparative models, which proves that this paper's algorithm has better recognition performance.

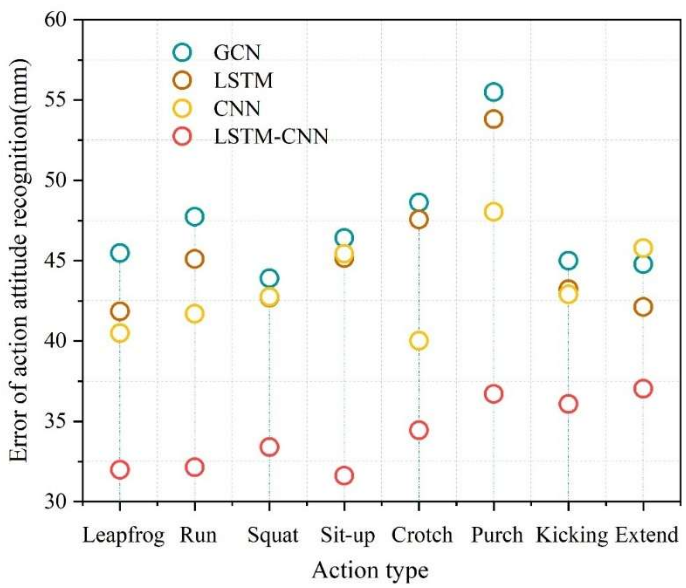
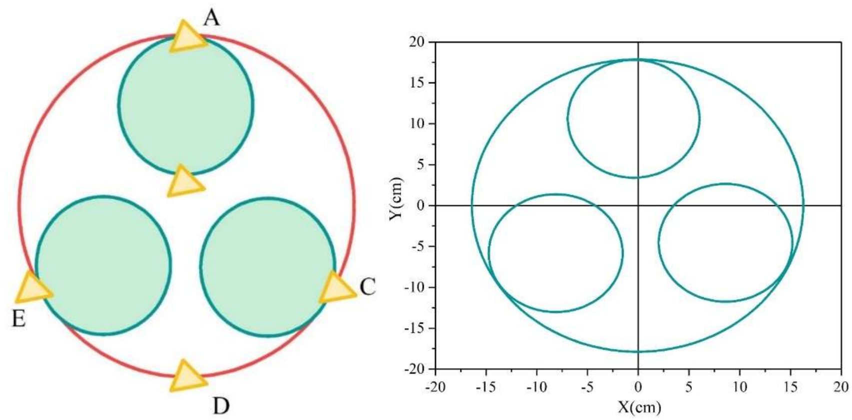


Figure 6: The actions of different models were compared to the test results

III. C. 2) Sports movement action recognition accuracy analysis

To verify whether the human motion recognition system can correctly locate the torso position and whether the error of joint point localization is within the acceptable range, the model is subjected to a three-circle route experiment. The results of the three-circle path and simulated walking schematic are shown in Fig. 7, where (a) represents the three-circle path schematic and (b) represents the motion trajectory simulation diagram. It can be seen that after the three-circle path walking recognition acquisition experiment for the recognition system, the improved recognition acquisition system can effectively capture the human plane motion trajectory. The test person is specified to walk in the order of five checkpoints and timed. It can be seen that the action movement recognition acquisition system simulates the body movement trajectory of the test person in good condition, the overall path planning record is complete, and the actual walking range is consistent with the prescribed walking range. Overall, the model is effective in recognizing the movements of different sports, and the structural advantages of the model enable it to more accurately capture the time series and the spatio-temporal relationship between joints, which improves the accuracy of movement recognition.



(a)Three circle diagram (b)Motion trajectory simulation

Figure 7: Three ring paths and analog walking diagrams

IV. Conclusion

In this study, under the guidance of the concept of big ideology education, a CNN-LSTM hybrid model based on sports action recognition is constructed, and the application effect of this model in the project-based teaching mode of sports courses in higher vocational colleges is investigated.

(1) After experiments, it is determined that the watermark intensity coefficient of this paper's method is set to 0.75; when the CNN-LSTM model is iterated for 40 times, its function loss value and recognition precision are finally stabilized at about 0.1251 and 0.9799. The accuracy of the four performance indexes, such as the precision rate and the recall rate of the CNN-LSTM hybrid model, reaches more than 99.9%, which proves that this paper's model has an application the effect is extremely good. It has the lowest error value for motion capture (between 0.1584mm-0.4128mm) and the highest accuracy rate of 99.08% for motion recognition.

(2) The results of empirical analysis show that the teaching model has an error of between 31.62mm-37.02mm in realizing the assessment of movement standardization in taijiquan demonstration, which is reduced by more than 20% compared with the comparison model; and it is able to more accurately capture the time series and the spatio-temporal relationship between the joints to improve the accuracy of movement recognition.

Funding

This work was supported by the Humanities and Social Sciences Research Project of Henan Province Universities [Grant Number 2024-ZZJH-224, Project: "Research on the Inheritance and Innovative Development of Central Plains Traditional Chinese Medicine Culture from the Perspective of Regional History"] and the Doctoral Research Startup Foundation of Henan Open University.

References

[1] Yu, B. (2024). Reform of Physical Education in Vocational Colleges under the Context of New Productive Forces. Journal of Modern Business and Economics, 1(2).
[2] Wang, H. (2023). Research on Physical Education Teaching in Higher Vocational Colleges Based on Physical Literacy and Professional Ability. Journal of Humanities, Arts and Social Science, 7(11).
[3] Shang, Y., & Shang, Y. (2022). RESEARCH ON THE INFLUENCE OF PHYSICAL EDUCATION IN HIGHER VOCATIONAL COLLEGES ON THE CULTIVATION OF STUDENTS'PSYCHOLOGICAL QUALITY. Psychiatria Danubina, 34(suppl 2), 196-196.
[4] Hongpeng, H. (2023). Analysis on Training Strategies of Students' Self-exercise Consciousness in Physical Education Teaching in Higher Vocational Colleges. Frontiers in Educational Research, 6(30).
[5] Martinez, C. (2022). Developing 21st century teaching skills: A case study of teaching and learning through project-based curriculum. Cogent Education, 9(1), 2024936.
[6] Markula, A., & Aksela, M. (2022). The key characteristics of project-based learning: how teachers implement projects in K-12 science education. Disciplinary and Interdisciplinary Science Education Research, 4(1), 2.
[7] Pan, G., Shankararaman, V., Koh, K., & Gan, S. (2021). Students' evaluation of teaching in the project-based learning programme: An instrument and a development process. The International Journal of Management Education, 19(2), 100501.

[8] Simonton, K. L., Layne, T. E., & Irwin, C. C. (2021). Project-based learning and its potential in physical education: an instructional model inquiry. Curriculum Studies in Health and Physical Education, 12(1), 36-52.
[9] Calderón, A., Scanlon, D., MacPhail, A., & Moody, B. (2021). An integrated blended learning approach for physical education teacher education programmes: teacher educators' and pre-service teachers' experiences. Physical Education and Sport Pedagogy, 26(6), 562-577.
[10] Mulya, G. (2023). The Effectiveness of Project Based Learning in Improving Student Learning Results and Motivation in Physical Education Learning. Indonesian journal of sport management, 3(2), 273-279.
[11] Ginanjar, S., Resmana, D., & Anugrah, S. M. (2024). Comparing project-based learning with conventional models: Enhancing students' enjoyment of physical education. Edu Sportivo: Indonesian Journal of Physical Education, 5(1), 64-81.
[12] Wen, P., & Wang, M. (2020). Project-based Teaching Model of Physical Education Based on OBE. International Journal of New Developments in Education, 2(7), 32-35.
[13] Sgro, F., Quinto, A., Platania, F., & Lipoma, M. (2019). Assessing the impact of a physical education project based on games approach on the actual motor competence of primary school children. Journal of Physical Education and Sport, 19, 781-786.
[14] Bafadal, M. F., & Haetami, M. (2021). The Role of Project-Based Learning (PjBL) Models in Improving the Ability to Organize Wrestling Sports Competitions for Physical Education Students FKIP Tanjungpura University. Kinestetik: Jurnal Ilmiah Pendidikan Jasmani, 5(3), 642-648.
[15] Song, Y. (2023). Research On Ideological and Political Implementation Path of Higher Vocational Physical Education Curriculum. International Journal of Vocational and Technical Education, 3(1), 59.
[16] Li, R., & Meng, X. (2023). Practical Exploration of Public Physical Education Courses in Higher Vocational Education from the Perspective of Curriculum Ideology and Politics—Taking Tennis Course as an Example. Advances in Physical Education, 13(4), 244-254.
[17] Goncharuk, C. V., Zhilina, L. V., & Mironova, T. A. (2022). Management of physical culture and health-improving activities in the higher education institution based on the project approach. Theory and Practice of Physical Culture, (5), 51-53.
[18] Hamzah, N. A. P., Rusli, K., & Janwar, M. (2025). HOLISTIC LEARNING IN PHYSICAL EDUCATION WITH PROJECT-BASED LEARNING: INSIGHTS FROM A STUDY ON BADMINTON TECHNIQUES. Journal Physical Education and Outdoor Activity, 1(1), 30-45.
[19] Lu, H. F. (2023). Statistical learning in sports education: A case study on improving quantitative analysis skills through project-based learning. Journal of Hospitality, Leisure, Sport & Tourism Education, 32, 100417.
[20] Yara Srinivas & Avatharam Ganivada. (2024). A modified inter-frame difference method for detection of moving objects in videos. International Journal of Information Technology,17(2),1-6.
[21] Sheng Hu,Gongjin Yuan,Kaifeng Hu,Cong Liu & Minghu Wu. (2023). Non-Intrusive Load Identification Method Based on KPCA-IGWO-RF. Energies,16(12),
[22] Agnieszka Duraj,Piotr S. Szczepaniak & Artur Sadok. (2025). Detection of Anomalies in Data Streams Using the LSTM-CNN Model. Sensors,25(5),1610-1610.
[23] C. Selvan,R. Senthil Kumar,S. Iwin Thanakumar Joseph,P. Malin Bruntha,M. Amanullah & V. Arulkumar. (2025). Traffic Prediction Using GPS Based Cloud Data Through RNN-LSTM-CNN Models: Addressing Road Congestion, Safety, and Sustainability in Smart Cities. SN Computer Science,6(2),159-159.