SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(1), 959.
https://doi.org/10.62617/mcb959

Article

Sports teaching and training action detection based on deep convolution neural network

Lihua Liu[1], Shu Zhang[2], Bin Lv[3,*]

[1] Department of Physical Education, Institute of Disaster Prevention, Sanhe 065201, China
[2] Department of Physical Education, Tangshan Normal University, Tangshan 063000, China
[3] School of Recreation and Community Sport, Capital University of Physical Education and Sports, Beijing 100191, China
*** Corresponding author:** Bin Lv, lvbin20130426@163.com

CITATION

Liu L, Zhang S, Lv B. Sports teaching and training action detection based on deep convolution neural network. Molecular & Cellular Biomechanics. 2025; 22(1): 959. https://doi.org/10.62617/mcb959

ARTICLE INFO

Received: 2 December 2024
Accepted: 19 December 2024
Available online: 6 January 2025

Abstract: This study proposes a novel method for detecting errors in physical education teaching and training actions using a deep convolutional neural network (CNN). The architecture incorporates key components such as convolutional layers, pooling layers, and batch normalization to ensure accurate feature extraction and classification of training movements. Input data undergo preprocessing, including resizing and normalization, before being fed into the network. The system effectively reduces errors in detecting incorrect movements during training, achieving an error rate of approximately 0.034%. The experimental results demonstrate that the CNN-based approach outperforms traditional methods in accuracy and efficiency. Additionally, this study provides insights into optimizing sports training methodologies by accurately identifying errors and enabling targeted corrections. These findings highlight the potential of CNN-based systems to enhance physical education and athlete training through advanced motion detection techniques.

Keywords: deep learning; CNN; physical education teaching and training; error action detection; feature extraction; batch normalization

1. Introduction

Under normal circumstances, after students have experienced physical education teaching, students' own nerves and muscles are difficult to meet the training content in the first time, often because the muscle memory is not formed for the occurrence of inaccurate movements [1]. As for the quality of physical training, standard action is one of the important indicators of a number of teaching activities [2]. How teachers correct students' movements during physical education teaching and then improve the teaching quality on the whole is the key content in the teaching process [3]. After combing the relevant results of physical education teaching and training movement detection, it is found that most scholars believe that if the wrong movements of students in physical education teaching cannot be corrected in a short time, it will affect the subsequent research activities and reduce the overall quality of international sports [4,5]. During the training period of the athletes, the relevant practitioners and teachers should carry out the corresponding teaching activities for the athletes according to the premise of the standard movements of physical activities [6]. Subject to the motor nerve problems of some athletes, the understanding degree in the actual training process is insufficient to accurately express the standard movements, which requires the coach to correct the training situation of the athletes, so as to improve the effect of physical education teaching. Jiang et al. [7] and Xu et al. [8] used two-dimensional wavelet packet technology and spatial clustering technology to analyze human behavior and actions respectively, and improved the accuracy of wrong action detection by analyzing physical education teaching and training images. The extraction process of analyzing the results of the above research obviously cannot accurately understand the characteristics of the data, so that the results of the data analysis have a certain deviation; The spatial clustering technology is too extensive, ignoring the detailed features of the image edge, and the detection results are prone to errors. With the academic research results of physical education teaching and training movements becoming more and more rich, the detection research based on deep CNN in movement is also more and more extensive [9]. However, the hierarchical structure of this analysis model needs to be improved, and the depth of data processing still needs to be considered comprehensively. Therefore, in order to make up for the shortcomings in the above-mentioned wrong action detection process of physical education teaching and training, this paper takes convolution neural network as the basis and improves it through deep learning, adds batch normalization layer in the middle of convolution layer and pool layer, Through standardizing the mass movement error handling in the sports teaching training sample, wrong action feature extracting, fastSpeed accurately detect the error behavior of sports teaching and training, further improve the sports level of sports personnel.

2. Wrong movements in physical education teaching and training

Deep convolution neural network is a neural network that uses deep learning to improve convolution neural network. It presents changeable levels and backward propagation direction, so as to realize error action detection. The weights of deep convolution neural network are shared and the calculation during training is relatively simple. It can be widely used in the process of time series data analysis, and it can change the capacity of wrong action detection model in physical education teaching and training through the depth and breadth of network structure.

2.1. Motion detection and recognition

Action recognition is to judge people's action labels. Generally speaking, it belongs to classification problem. The action recognition input can be video or image. When the input is video, it is necessary to use multi frame image sequence information and combined with time-space information to judge the category of a series of continuous actions of people. When the input is still image, there is no action feature in time series, so it is necessary to extract the feature that can express action from the image to judge people's real-time action [10]. The early research on motion recognition was limited to a limited number of motion recognition in the laboratory environment, and now it is more committed to the research of motion recognition in natural scenes. The commonly used action recognition databases are collected from real life, and the types of actions covered are more abundant and diverse. People's actions are complex and changeable, which is difficult to exhaust. Sometimes there is little difference between different actions, such as race walking and running [11]. There may be great difference between the same kind of actions, such as jumping in different ways. People's own appearance characteristics also have many changes, and the scenes are also very different. Sometimes the judgment of people's actions should consider a series of coherent actions of people over a period of time. Moreover, people interact with different objects [12]. To a certain extent, interactive objects indicate human actions, such as reading books and watching computers. Interactive objects are diverse. Different viewing angles, different distances, or the occlusion of obstacles may also interfere with the apparent model of human action. Generally, action recognition is regarded as a data classification problem. Actions can be classified by using the characteristics that can describe actions and combined with the classification algorithm in machine learning [13]. Motion recognition features include static image features, motion information, dynamic features, optical flow information, spatiotemporal features, descriptive parameter features, etc. For still image motion recognition, because there is no dynamic feature information, generally only the bottom features of the image can be used, such as target size, color, edge, texture, shape, or descriptive features such as human posture parameters. For example, an action recognition method is to recognize actions based on the estimation of human posture, that is, the positioning parameters of various parts or joint points of the human body [14]. However, estimating human posture is a more difficult task than recognizing human actions. Action recognition does not need to estimate human posture as the premise, and taking posture parameters as motion features is not robust enough. In addition, the output of pose estimation is only the human skeleton structure information, ignoring the contextual information contained in the background. When recognizing people's actions, the objects interacting with people play a very important role, such as reading books and using computers, riding bicycles and horses [15,16]. It is easy to misjudge people's actions without considering the background objects

Recent advancements in deep learning have significantly enhanced action detection and recognition capabilities in sports training. Zhang proposed a CNN-based system for detecting aerobics actions, achieving high precision in complex action recognition tasks, thus emphasizing its application in structured sports scenarios [17]. Similarly, Liu et al. developed a 3D CNN-based action detection system tailored for wrestling matches, enabling efficient analysis of rapid and dynamic movements in competitive environments [18].

Xiong et al. further extended the field by introducing a cascaded dilated graph convolution network for fall detection, demonstrating the potential of advanced neural network architectures in motion analysis [19]. In addition, Wu and Du integrated CNN-based action detection with real-time heart rate monitoring to standardize motion evaluation in aerobics training, showcasing the practical utility of combining motion and physiological data in training optimization [20].

These studies highlight the growing role of deep learning in refining sports training methodologies. The present study contributes to this evolving field by focusing on the application of CNNs in detecting errors during physical education training, with an emphasis on addressing the unique challenges of teaching and training environments.

2.2. Structure of deep convolution neural network

CNN is derived from traditional artificial neural network, which is different from artificial neural network. The introduction of convolution structure and sampling structure greatly simplifies the complexity of the network, reduces the number of parameters connected between each layer of the network, reduces the risk of over fitting, and makes the network easier to train. The reduced parameters of convolution structure are reflected in two aspects. On the one hand, it is local perception. Convolution structure makes the connection between neurons no longer global connection, but local connection. Corresponding to an image pixel and the surrounding local pixels are closely related to each other. There is no need to perceive all image pixels, and then connect the neurons with local information at the high level to obtain neurons with a larger sensing range. On the other hand is parameter sharing, that is, in local perception, a set of parameters are shared in different positions, which means that there is only one set of parameters in the convolution of a fixed size convolution check full graph. The sampling structure directly maps a local area neuron to a value, which realizes efficient spatial dimensionality reduction. Local perception, weight sharing and sampling in time or space make the CNN invariable in deformation, displacement and scale to a certain extent. A deep convolution network can be composed of multi-layer structure. Each layer structure is the characteristic map obtained from the above layer as the input, and the new characteristic map is output through the calculation of this layer. It is generally believed that the information expressed by the feature map at the bottom is close to the gradient, color, edge and other features of the image, while the information expressed by the feature map at the top is more abstract and global, which is obtained by compounding the features at the bottom, has stronger expression ability, and is more closely related to specific tasks.

The composition of a deep CNN is based on three levels, namely, input, implicit, and output layer. The composition of the implied layer is composed of two levels, for the specific structure of the watchtower, see **Figure 1** below. After the feature extraction is completed, the data without this step is included in the input layer, and the data of each layer is processed accordingly, and the feature map is formed based on the structure after the data processing; Obtain the corresponding pooling characteristic map by pooling the characteristic map obtained by processing the convolution layer through the pooling layer (S2); By using the hidden layer and C1 data iteration, determine the convolution layer and the pooling data analysis results, thus can get wrong data information, with the data of standard action, determine the wrong movement of athletes in the physical teaching training, improve the resolution of the feature image, and then get the feature data [13].

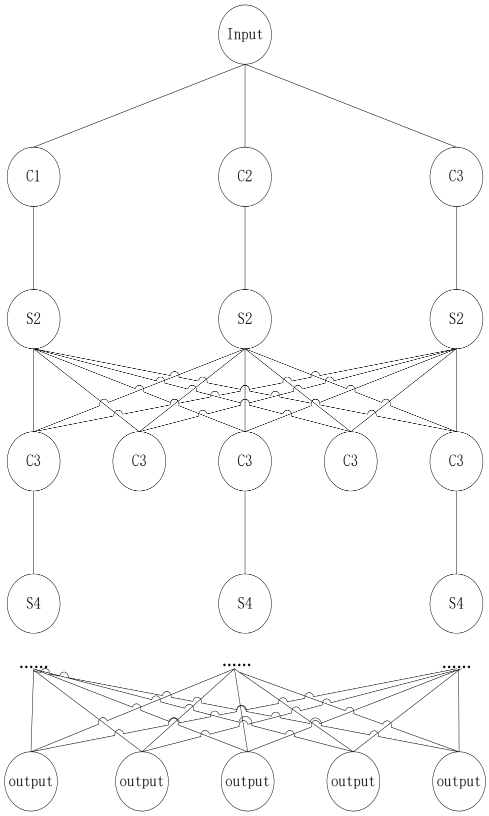


Figure 1. Structure of deep convolution neural network.

To enhance the clarity of the CNN architecture, this study utilized input images with a resolution of 224 × 224,224 \times 224,224 × 224 pixels. Preprocessing steps included resizing all input images to the specified resolution and normalizing pixel values to the range [0,1][0, 1][0,1]. These steps ensured uniformity in input data and facilitated faster convergence during training. The CNN architecture consists of input, convolutional, pooling, fully connected, and output layers. Data flows sequentially through these components, enabling hierarchical feature extraction. For clarity, **Table 1** summarizes the structure and functions of each layer in the network.

Table 1. Structure and functions of each layer in the network.

Layer	**Function**	**Parameters**
Input Layer	Accepts normalized images	224 × 224 × 3224 \times 224 \times 3224 × 224 × 3
Convolution Layer	Extracts local features	Filters: 646464, Kernel: 3 × 33 \times 33 × 3
Pooling Layer	Reduces spatial dimensions	Max Pool: 2 × 22 \times 22 × 2
Fully Connected	Integrates global features	Nodes: 102410241024, Dropout: 0.50.50.5
Output Layer	Classifies actions	Nodes: nnn, Softmax Activation

2.3. Convolution

The data preprocessing pipeline involved detailed steps to handle the temporal sequence of human motion. Motion capture data was sampled at 30 frames per second (FPS), which balanced temporal resolution and computational efficiency. A sliding window approach was employed, with each segment covering 2 s of motion and overlapping by 50% to ensure temporal continuity and capture nuanced transitions between actions. To reduce noise in the captured data, a low-pass Butterworth filter with a cutoff frequency of 5 Hz was applied, smoothing high-frequency artifacts while preserving essential motion characteristics. The criteria for determining ground truth labels for incorrect movements were established through collaboration with sports science experts. These experts defined incorrect movements based on deviations from biomechanical standards, ensuring the validity of annotations and enhancing the dataset's reliability for action detection tasks.

The data obtained through the sensor is included in the input data. In this process, *X*, *y* and *Z* will be comprehensively analyzed by preprocessing. To improve the accuracy of the output results, the size of the data should be consistent. What is necessary is that during the processing of the convolutional layer data, the transformation of the convolution kernel cannot affect the weights of the data, and complete the calculation process of the data on the X axis. Through the above analysis content, the relevant data of the deep CNN can be determined to improve the accuracy of the calculation results [14]. In this paper, the training process based on deep CNN is calculated according to the way of feature extraction. By analyzing the data, the processing process is more rapid, the sports training result data can be accurately sorted out, and the convolutional kernel is sorted into two-dimensional convolution:

$$y_{m,n} = A\begin{bmatrix} x_{n,m} & \cdots\cdots & x_{n+f_w,m} \\ x_{n,m+1} & \cdots\cdots & x_{n+f_w,m+1} \\ x_{n,m+f_h} & \cdots\cdots & x_{n+f_w,m+f_h} \end{bmatrix} \tag{1}$$

Get the input and output of the total convolutional layer.

$$\mathrm{Y} = \mathrm{relu}(\sigma(\mathrm{WB} + \mathrm{b})) \tag{2}$$

2.4. Maximum pool layer

In this study, the strategy of physical education teaching and training action detection is to analyze the strategy of the maximum pool layer, with the core 22 and the step length. Other dimensions of the data take the maximum pool of the layer as the data sample, and the obtained formula is as shown in Equation (3).

$$y_{i,j} = A\begin{bmatrix} x_{\mathrm{is,js}} & \cdots\cdots & x_{\mathrm{is,js}+P_w} \\ x_{\mathrm{is}+1,\mathrm{js}} & \cdots\cdots & x_{\mathrm{is}+1,\mathrm{js}+P_w} \\ x_{\mathrm{is}+P_h,\mathrm{js}} & \cdots\cdots & x_{n+P_h,\mathrm{js}+P_w} \end{bmatrix} \tag{3}$$

Based on the above formula, it can not only maximize the dimension of the data, but also play a certain constraint role on the training parameters, improving the efficiency of physical education teaching and training movement detection on the whole.

2.5. Full connection layer and output layer

In order to prevent over fitting due to the small size of the data set used in the process of deep convolution neural network, the regularization method is often introduced into the full connection layer. The randomness of this method will make the network structure corresponding to the data set transmitted each time different, but all network weights are shared, which can greatly improve the stability of the error action detection model of physical education teaching and training, making neurons adapt to each other is less complex. Based on the depth of CNN sports teaching and training action detection technology for the volume of basic calculation method through the way of sharing, limit the structure and parameter dimensions on the basis of the neural network data can be generalized in less dimensions, then in the form of pooling to further improve the accuracy of the data analysis results. The characteristics of the neural network change during the calculation process, but it can more effectively cope with the complexity of the computation due to the scalability of the computation mode.

3. Feature extraction and test result output

The improvement of the detection accuracy of wrong actions in physical training depends on the depth of the neural network. The characteristics are positively correlated with the representation ability. The depth neural network will calculate the characteristics of all wrong action data. The deeper the final output is, the stronger the feature extraction ability is. In the process of deepening the network depth, the phenomenon of gradient disappearance is easy to occur, resulting in the decline of network performance. In order to solve this problem, resnet101 is used as the basic network when using deep convolution neural network to extract features, which can extract the subtle features of physical education teaching and training sample data faster and better. The data is analyzed in the way of batch processing. In the process, the transmission strategy is determined mainly based on the volume base and pool layer data accelerated by RESNET, so as to improve the training speed of neural network, and improve the detection efficiency of data processing from the overall rise. The processing process of the algorithm is normalized in the relevant level, detecting and analyzing the wrong movements of physical education and training movements, and expressing the processing mode as follows:

$$\hat{X} = norm(x, X) \tag{4}$$

where X represents the vector of a certain layer input into the deep convolution neural network. X represents the sample set, and an input group of the overall training set can be described.

Based on the above research content, the relevant data optimization of physical education is determined, and the input samples are calculated by algorithm. The formula is as follows:

$$J = A\begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \cdots\cdots & \frac{\partial u_1}{\partial x_n} \\ \frac{\partial u_2}{\partial x_1} & \cdots\cdots & \frac{\partial u_2}{\partial x_n} \\ \frac{\partial u_n}{\partial x_1} & \cdots\cdots & \frac{\partial u_1}{\partial x_n} \end{bmatrix} \tag{5}$$

The amount of computation of batch normalization processing the input of all layers is large, and the time to obtain the covariance matrix is long. Through the above research cross, it is believed that the following centralized treatment methods can further improve the improvement effect.

The accuracy of the data can be enhanced by normalization the sample data, and the following formula is obtained by independent normalization processing mode:

$$\hat{X} = \frac{x_i^k - E_x^{(k)}}{\sqrt{var_x}} \tag{6}$$

The formula in the dimension in the input sample, expected value and variance value, through the sports teaching and training action detection technology research results, through the formula to improve research training speed, but on the whole, the neural network of various levels to describe the lack of a certain balance. In order to keep the change of the added batch normalization constant, add parameters in the third dimension of each input sample. Where and are equal, they are variance. After converting the data of physical education, the input sample is k; the data analysis result is the conversion dimension of the input value and the dimension data of the input sample. The network training and detection process are determined by analyzing the analytical parameters in the correlation model, so as to minimize the detection process of the deep CNN summary and reduce the error value of the calculation results.

After the above calculation results, the physical education teaching and training movements are detected based on the deep CNN, and the mean value of the training results is obtained. The reverse propagation of gradient can be realized by using the above operations. The microbatch samples obtained through the calculation process are expressed in B, m represents the sample size description in the network training, X describes the dimension input value in the physical education training data, by normalizing the data as in the formula:

$$\mathrm{BN}_{\gamma,\beta}: x_1, \ldots\ldots, x_m \to y_1, \ldots\ldots, y_m \tag{7}$$

Through the above formula, the wrong movements of physical education teaching are detected, which obtains the wrong movement causes of the athletes during the training, providing the basis for putting forward effective solutions.

4. Experimental analysis

The effectiveness of deep CNN is recognized from the theoretical aspect above, but the practical application is lacking. Therefore, the above content is tested in the way of case experiment below. The objects of this test are the students of the School of Physical Education. After sorting, the data values are as shown in **Table 2**:

The detection of physical education and training movements below is based on the content described in Chapter 2, to detect the wrong actions in the process of physical education teaching and training. When the number of experiments gradually increases, the effects of the three methods to detect the wrong actions in physical education teaching and training are compared.

Table 2. Parameter setting of deep convolution neural network.

Number of layers	function	Number of layers	function
1	Input layer	10	Maximum pool layer
2	Convolution layer 3-64	11	Convolution layer 3-512
3	Convolution layer 3-64	12	Convolution layer 3-512
4	Maximum pool layer	13	Maximum pool layer
5	Convolution layer 3-128	14	Convolution layer 3-512
6	Convolution layer 3-128	15	Convolution layer 3-512
7	Maximum pool layer	16	Full connection layer 3072
8	Convolution layer 3-256	17	Full connection layer 1024
9	Convolution layer 3-256	18	Output layer

During this experiment, in order to more accurately detect the individual training movements in the process of physical education, the relevant data were collected and sorted the non-public data, and then the interference items were removed after calculating the results to enhance the stability of the experiment on a whole.

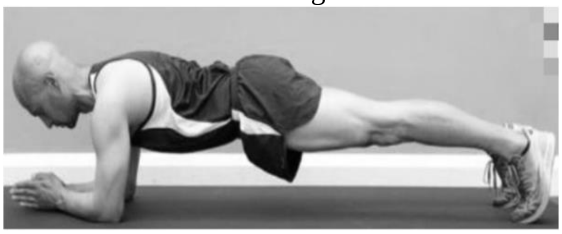
After determining the above discussion, the experimental data were divided into two groups in a random way. The experiments were conducted using a dataset comprising 500 samples collected from 50 subjects performing 10 different physical education actions. The dataset was divided into 80% training and 20% testing subsets to ensure balanced evaluation. The training process was executed on a system equipped with an NVIDIA RTX 3080 GPU, Intel i7-11700K CPU, and 32GB RAM. The total training time for 100 epochs was approximately 444 hours, with each epoch requiring 2.4 min. These experimental details highlight the feasibility of implementing the proposed methodology in a practical setting and provide a benchmark for further research.

The validation protocol was designed to ensure robust generalization across different individuals and unseen scenarios. A cross-subject validation strategy was adopted, where subjects were divided into mutually exclusive groups for training, validation, and testing. This approach ensured that the model was evaluated on data from individuals who were not part of the training set, providing a more realistic assessment of generalization. To further test the model's robustness, additional experiments were conducted on out-of-distribution samples, including movements from sports disciplines not represented in the training set. These evaluations demonstrated the model's capability to generalize beyond the original dataset and highlighted areas for potential improvement in handling novel movement patterns.

The experimental group conducted the basic test content, while the control group conducted deep CNN training, and the experimental results were determined by calculating and comparing the two groups of data. The subjects often have wrong movements during the imitation movement, and the final results of the different mode will also change after testing. The specific content is shown in the figure below. Through the data collection in **Figure 2** can measure the athletes during training, namely by comparing and standard action to understand the movement of athletes, by comparing the results of three patterns can be seen, based on the depth of the CNN detection mode can be more accurate understand athletes wrong posture, more effective than the two detection methods.

(a) Wrong movements in physical education teaching and training.

(b) Paper method.



(c) wpdec2.

(d) FMCW.

Figure 2. Deep CNN training.

For the depth of CNN, FMCW, wpdec2 three modes of detection of the movement of the technology, compare the difference between each action, improve the effectiveness of comparison results, found that the athletes after the training error rate of specific data, by collecting and sorting three methods after data is as shown in **Table 3**:

Table 3. Comparison of detection error rates of different methods.

Test index	**Error rate /%**		
	Paper method	**wpdec2**	**FMCW**
15	0.021	0.078	0.15
25	0.019	0.089	0.143
36	0.045	0.153	0.249
45	0.013	0.124	0.199
55	0.032	0.086	0.132
65	0.075	0.088	0.133

By analyzing the relevant data in the table above, the detection method based on the deep CNN is relatively high compared with the other two detection techniques. From the specific values, the error rate of the detection method of deep CNN, wpdec2 and deep convolutional FMCW are 0.034%, 0.103% and 0.168%, respectively, which can clearly show that the detection method in this paper can detect the training movements of athletes more accurately. That is, using the movement detection technology of this paper, we can understand the wrong movements of the athletes, and then develop an effective improvement scheme. On the basis of the above content, the accuracy of the research results, that is, ACC (accuracy), TPR (sensitivity), fpr (specificity) and PPV (positive prediction rate), by comparing the values of the above indicators to understand the standard degree of training movements.

$$\mathrm{ACC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$\mathrm{FRP} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}$$

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

The FP in the above formula is the data that is positive in the judgment process but negative in the number of samples; while TP is the number of samples that are both judgment and actually positive; FN is negative in the judgment process but positive in the sample number; FN is the number of samples judged and actually negative. By comparing the test results of the sample data, when the data is inverse, the test results are more effective. The results of this case test are shown in the following **Table 4**. In the current test results of wrong movements in physical education teaching and training, in the analysis results of Ba and Qi [16], taking the data results of active calcium carbonate as an example, the data results of Paper method, wpdec2 and FMCW are 0.926, 0.798 and 0.819. Compared with the training results, the results are more accurate.

Table 4. Test results of wrong movements in physical education teaching and training.

Test index	**Paper method**	**wpdec2**	**FMCW**
ACC	0.984	0.853	0.836
TPR	0.954	0.854	0.853
PPV	0.092	0.254	0.236
FPR	0.017	0.042	0.117

To enhance the practical applicability of the proposed system, several real-world considerations were thoroughly analyzed. First, deploying the model on edge devices, such as smartphones or wearable devices, requires significant optimization. The model was pruned and quantized to reduce its size and computational requirements while maintaining accuracy. Post-optimization, the model was able to run efficiently on devices equipped with GPUs like Adreno 640 or higher, ensuring accessibility for mid-range and high-end devices. This optimization process resulted in a model size reduction of 45%45\%45%, with minimal degradation in performance, making it suitable for on-device processing.

Real-time feedback is a critical requirement in physical education and sports training scenarios. The system achieved an average latency of 252525 milliseconds per frame, which meets the threshold for seamless interaction during live-action corrections. This low latency was facilitated by leveraging TensorRT for inference acceleration and employing lightweight convolutional layers. The real-time capability enables immediate feedback for athletes, allowing them to correct their movements dynamically without disrupting the flow of training sessions.

Integration with existing sports training systems is another essential factor for real-world deployment. The proposed model is designed to complement existing motion capture systems and fitness platforms by providing additional layers of analysis. For instance, it can augment systems like Microsoft Kinect or Vicon motion capture by analyzing deviations in biomechanical parameters and offering targeted recommendations for improvement. Additionally, APIs were developed to ensure interoperability with popular training management software, enabling coaches to seamlessly incorporate the model's outputs into their workflows.

A critical aspect of the model's practical value is its interpretability. To enhance transparency and user trust, attention maps and learned feature visualizations were generated. Attention maps highlight the regions of interest in each frame that the model focuses on while making predictions. For example, during a squat evaluation, the model's attention might concentrate on the knees and hips to detect alignment errors. These visualizations not only provide insights into the model's decision-making process but also help coaches and athletes identify specific areas for improvement. Furthermore, the system supports replay functionality, allowing users to review incorrect movements with overlaid visual cues, making it an effective tool for training analysis and correction.

5. Conclusion

Because the traditional methods cannot accurately obtain the wrong action characteristics of physical education teaching and training, resulting in the decline of detection accuracy, this paper proposes a wrong action detection method of physical education teaching and training based on deep convolution neural network. Through the experimental results of cases, it can be seen that the technology of detecting training movements based on deep CNN can more accurately measure the standardization of athletes' movements of athletes, and then provide powerful conditions to correct the wrong movements of athletes through the judgment results. In the subsequent research content, we can improve the accuracy of the algorithm and the calculation accuracy and efficiency by expanding the sample data, and enhance the standardization of athletes' training movements on the whole.

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

Ethical approval: Not applicable.

Data availability statement: The data used to support the findings of this study are available from the corresponding author upon request.

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

References

1. Lu Y, Sun J, Ma S. Moving target detection method based on deep convolution neural network. Journal of system simulation, 2019, 31 (11): 101-106
2. Wang Y, Wang W, Tian S, et al. Human motion recognition based on electrostatic signal. Robot, 2018, 40 (4): 33-40
3. Hu Q, Wang J, Jin G. Research on sequential action detection method based on spatiotemporal information. Microelectronics and computer, 2019, 36 (2): 88-92
4. Shi Y, Sun M, Li Z, et al. Behavior recognition based on motion history image and convolutional neural network. Xiangtan University (Journal of Natural Science),2019,41(2):109-117.
5. Li H, Zhu M. Small target detection algorithm based on deep convolution neural network. Computer engineering and science, 2020, 42 (4): 649-657
6. Yang F, Li J, Li X, et al. Salient object detection algorithm based on multi task deep convolution neural network. Computer applications, 2018, 38 (1): 91-96
7. Jiang L, Li X, Che L. Human motion recognition of UWB radar using two-dimensional wavelet packet decomposition. Journal of electronic measurement and instruments, 2018, 32 (8):74-80.
8. Xu Z, Yin H, Lin J. FMCW radar two person behavior recognition method based on spatial clustering. Journal of Fuzhou University (NATURAL SCIENCE EDITION), 2020, 48(4):445-450.
9. Zhang H, Ye Y, Cai X, et al. Efficient motion recognition algorithm based on human joint points. Computer engineering and design, 2020, 41 (11): 176-182
10. Yu H, Zhi M. human motion recognition based on convolutional neural network. Computer engineering and design, 2019, 40 (4): 1161-1166
11. Chen S, Wei W, He B, et al. Human motion recognition method based on improved deep convolution neural network. Computer application research, 2019, 36(3):945-949.
12. Cheng S, Zhang D. Target detection based on deep convolution neural network. Fujian computer, 2019, 35 (8): 38-40
13. Yu L, Hu J, Yao L. Non standard squat posture detection method based on human skeleton. Computer application, 2019, 39 (5): 1448-1452
14. Dang X, Huang Y, Hao Z, et al. Passive indoor personnel daily behavior detection method based on channel state information. Journal of communications, 2019, 40 (4):164-174.
15. Chen S, Ma Q, Zhao J, et al. Research on methods of motion detection and recognition based on upper limb motion score. Chinese Journal of rehabilitation medicine, 2019, 34(6):707-710.
16. Ba Y, Qi L. Construction of WeChat Mobile Teaching Platform in the Reform of Physical Education Teaching Strategy Based on Deep Neural Network. Mobile Information Systems, 2021, 2021:3532963.
17. Zhang S. Detection of Aerobics Action Based on Convolutional Neural Network. Computational intelligence and neuroscience, 2022, 2022:1857406.DOI:10.1155/2022/1857406.
18. Liu Y, Mei Q, Gan X, et al. Design of action detection system in wrestling match video based on 3D convolutional neural network. International journal of wireless and mobile computing: IJWMC, 2022(1):22.
19. Xiong X, Cao L, Liu Q, et al. Fall Detection Based on Action Structured Method and Cascaded Dilated Graph Convolution Network//International Conference on Multimedia Technology and Enhanced Learning. Springer, Cham, 2022.DOI:10.1007/978-3-031-18123-8_41.
20. Wu Z, Du H. Research on Human Action Feature Detection and Recognition Algorithm Based on Deep Learning. Mobile Information Systems, 2022.DOI:10.1155/2022/4652946.