SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2024, 21(3), 490.
https://doi.org/10.62617/mcb490

Article

Analysis of athletes' technical action based on deep learning

Xiaoning Zhang

Department of the Sports, Xi'an International Studies University, Xi'an 710128, China; zhangxiao202401@126.com

CITATION

Zhang X. Analysis of athletes' technical action based on deep learning. Molecular & Cellular Biomechanics. 2024; 21(3): 490. https://doi.org/10.62617/mcb490

ARTICLE INFO

Received: 9 October 2024
Accepted: 25 October 2024
Available online: 26 November 2024

COPYRIGHT

Copyright © 2024 by author(s). *Molecular & Cellular Biomechanics* is published by Sin-Chn Scientific Press Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license. https://creativecommons.org/licenses/by/4.0/

Abstract: In order to raise athletes' technical proficiency and overall performance, this paper uses deep learning technology to precisely assess table tennis technical activities. The paper builds a hybrid neural network model for technical action analysis of table tennis players, based on Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) in deep learning. First, CNN extracts the spatial characteristics from the video frames. After that, an LSTM processes these data in a time series to accurately recognize and classify the movements of athletes. Lastly, through trials, the model is trained and evaluated using the Table Tennis Tracking Network (TTNET) dataset. The experimental findings demonstrate that: (1) In table tennis technical action analysis, the suggested deep learning model performs better than other comparative models. The efficacy of combining convolution with sequence learning is fully demonstrated by the CNN-LSTM hybrid model, which performs best in all indicators when compared to the CNN, LSTM, Multi-Objective Function (MOF), and Convolutional Neural Network—Long Short-Term Memory (CNN-LSTM) combined models. Its accuracy rate, precision rate, recall rate, and F1 score are 0.923, 0.918, 0.925, and 0.921, respectively. In contrast, the performances of LSTM and CNN are also excellent, but the performance of MOF model is relatively low; (2) In the classification of technical actions, the model has the highest classification accuracy of service, reaching 0.930, and the classification accuracy of other technical actions such as forehand stroke, backhand stroke and spike is also above 0.9, and the time sequence consistency index also maintains a high level, indicating that the model can effectively identify and analyze table tennis technical actions. In addition, the performance evaluation of real-time feedback shows that the model can achieve low processing time and feedback delay in video data processing with different lengths, which ensures the real-time and reliability in practical applications. These results show that the proposed model can not only provide accurate technical action recognition, but also provide timely and effective feedback in practical application, which has high practical value. The results of this paper prove the potential of deep learning technology in the analysis of athletes' technical actions, and provide scientific basis and effective tools for the technical training and optimization of table tennis players.

Keywords: deep learning; table tennis player; CNN-LSTM; accuracy; technical action analysis

1. Introduction

Deep learning and computer vision technologies have demonstrated significant promise in the realm of sports analysis with the ongoing advancements in sports technology. The technical actions of the athletes are assessed using manual observation and video capture in the conventional sports analysis approach, which is time-consuming, labour-intensive, and susceptible to subjectivity [1]. Deep learning technology has advanced quickly in recent years, opening new possibilities for sports research. It can now automatically extract complex features from video data and precisely classify and evaluate athletes' technical actions.

Deep learning is a significant subfield of machine learning that originated with the creation of artificial neural networks [2]. The first neural network model was presented or pattern recognition and classification applications as early as the 1960s. Nevertheless, the performance of these early models in real-world applications is not optimal because of the limitations in processing power and the absence of data [3]. Deep learning has advanced significantly since the turn of the twenty-first century thanks to advancements in computer hardware, particularly the widespread use of Graphics Processing Units (GPUs). With the use of layered network structures, deep learning models—Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) in particular—can automatically extract and learn characteristics from vast amounts of data, improving prediction and classification accuracy [4]. In action analysis, deep learning technology greatly improves the efficiency and accuracy of analysis through automatic data processing and feature extraction. For example, the successful application of CNN in image processing makes it an ideal choice for processing video data, while the LSTM is excellent in processing time series data and can capture the time series characteristics of athletes' technical actions [5,6]. The combination of these technologies provides a powerful tool for sports analysis, which enables athletes' technical actions to be quickly and accurately identified and classified from massive video data.

This paper aims to systematically analyze the technical actions of table tennis players based on deep learning technology. This paper investigates the automatic extraction of technical action elements from table tennis films using a deep learning model, followed by time sequence analysis and classification. To fully utilize the benefits of deep learning in time series modeling and spatial feature extraction, a CNN/LSTM hybrid model is used. Through training and testing, this paper evaluates the accuracy of the model in identifying table tennis technical actions, and analyzes the classification effect and time sequence consistency of different technical actions in detail.

2. Literature review

In recent years, video action analysis technology based on deep learning has been widely used in the field of sports. The following studies show the progress in this field. Khobdeh et al. proposed a method combining You Only Look Once (YOLO) and Long Short-Term Memory (LSTM) network to identify human behavior in basketball in complex environment. YOLO was used to detect players in frames, and LSTM was combined with fuzzy logic for final classification. This method had proved its superiority on several datasets, especially in dealing with complex background, blocked action and inconsistent lighting conditions [7]. The research in the field of table tennis action recognition has made remarkable progress in recent years, and many research teams have put forward innovative technical methods to improve the accuracy of action recognition and technical analysis. Song et al. proposed an action recognition framework based on You Only Look Once version 8 (YOLOv8)-Alphapose-two stream Spatio Temporal Graph Convolutional Network (2s-STGCN) to analyze the technical actions and tactical patterns in table tennis. Their study combined dynamic and static complex network analysis and community detection algorithm to evaluate the performance of 37 elite players in 8015 high-definition action videos in two self-built datasets (front player dataset and back player dataset). Experimental results showed that the performance of YOLOv8-Alphapose-2s-STGCN on these two datasets was better than other seven algorithms based on YOLOv8-Alphapose, such as Transformer, Bidirectional Gated Recurrent Unit (BiGRu) and so on. In addition, the study also evaluated the importance of the nodes (technologies) in the serve and catch network and the two-round (win or lose) network through various indicators, and analyzed the tactical style and mode by using the dynamic complex network to reveal the tactical similarities and differences between players and opponents [8]. Chalavadi and Yenduri proposed an adaptive time aggregation method for table tennis hitting action recognition, focusing on solving the challenges brought by high similarity between classes, occlusion and viewpoint change in fine-grained action recognition. They introduced a table tennis hitting 1.0 dataset, including 9000 unconstrained videos of 6 kinds of movements. This method used three-dimensional convolutional neural network (3D CNN) to extract features, and combined with time aggregation network to replace the traditional global average pooling layer to capture subtle time interaction. The introduced time aggregation network made use of the attention mechanism from the converter represented by the bidirectional encoder, which effectively improved the accuracy of action recognition, and its performance on the table tennis hitting 1.0 dataset was improved by 10% [9]. Zhang et al. proposed an algorithm based on variable-length input CNN by analyzing the trajectory of table tennis bat, and constructed a table tennis technical action dataset TTMD 6 for research. Different from the traditional method using multi-joint data, their study only recognized the action through the trajectory of the racket, which significantly sped up the recognition process. This trajectory-based identification method had a wide application potential in biomechanical analysis, auxiliary training system and other fields [10]. These studies showed that the application of deep learning technology in sports video analysis not only improved the accuracy and efficiency of action recognition, but also provided important support for coaches and athletes in training and competition decision-making.

3. Design of CNN-LSTM hybrid neural network model

In order to assess the table tennis video action dataset, this paper presents a hybrid model that combines CNN and LSTM [11]. According to the concept, CNN oversees using a sliding window and weight-sharing mechanism to extract local spatial characteristics from video sequences and forwards them to the following layer. In order to optimize each technology's advantages, LSTM combines time series processing with time series feature extraction from video data [12]. Furthermore, the use of deep learning models enables the automatic extraction of action features, which not only streamlines the formerly labour-intensive process of manually extracting features but also greatly increases the accuracy of action detection. In the task of processing time series information, especially identifying human activities and behaviors, LSTM has become an ideal network structure because it can learn long-term dependencies [13]. The hybrid model in this paper is based on this design concept, aiming at improving the efficiency and accuracy of table tennis technical actions recognition.

Therefore, this paper proposes a hybrid neural network model based on CNN feature construction strategy and LSTM end-to-end feature construction to identify basic table tennis technical actions. Firstly, the basic table tennis technical action video data is pre-processed to form a sequence of images, which are randomly divided into basic practice dataset and test dataset. While test datasets are used to validate the model's performance features, training data is utilized to confirm the model's design and parameter adjustments. Here, the spatial characteristics of each frame are obtained through CNN, and then the output of CNN network is transmitted to the LSTM network module after scale adjustment to obtain the time characteristics of the sequence. The average data output by LSTM at each time is calculated and finally passed to the classifier to predict the final classification result [14]. The framework of action recognition model based on CNN-LSTM is shown in **Figure 1**.

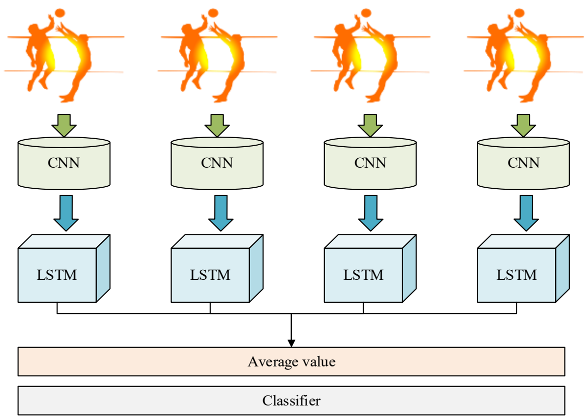

Figure 1. Framework of action recognition model based on CNN-LSTM.

Firstly, the model extracts the basic features of each frame image in the video sequence through CNN. Secondly, the CNN action representation value of each frame in the video is input into the LSTM unit respectively, which is used to learn signals such as actions and changes associated with video sequences. In the third step, the output value of each LSTM unit is linearly transformed, and the classification probability of the corresponding action behavior between the output of the classifier Softmax layer and the video sequence is determined by averaging these probabilities to determine its final action recognition accuracy. The design and function of each part of the network structure are introduced in detail below.

3.1. CNN processing layer

Firstly, the model obtains the spatial information of video frames through three-layer CNN, thus forming the representation characteristics of human actions. It is assumed that the maximum number of frames of the video frame data sequence input into the network is N [15]. The pixel matrix size of each image is P × Q, and the convolution operation adopts "padding = SAME", that is, when the sampling center position coincides with the edge, convolution calculation is started, and the size of the feature map after convolution remains unchanged to ensure that the size of the feature map does not change during forward propagation.

Let the convolution kernel size be $k \times k$, get the target features by sliding window method, and let v_{ij} be the pixel value of the pixel matrix at (i, j). Then all pixel values falling into the $(Q(i-1)+j)$th sliding window can define the window matrix X_{ij}, which is expressed by the following equation:

$$X_{ij} = \begin{pmatrix} v_{ij} & \cdots & v_{i(j+k-1)} \\ \vdots & \ddots & \vdots \\ v_{(i+k-1)j} & \cdots & v_{(i+k-1)(j+k-1)} \end{pmatrix} \tag{1}$$

For each window matrix, features can be extracted by convolution operation. A convolution operation is a mathematical procedure that involves sliding the convolution kernel (or filter) over each area of the window matrix. The result is the window features, which are obtained by adding the products of the kernel and the corresponding components of the window matrix. This equation can be used to express this process:

$$Y_{ij} = f(X_{ij} \otimes W + b) \tag{2}$$

Rectified Linear Unit (ReLU) function is widely used because of its simple calculation and fast convergence. The equation of ReLU function is as follows:

$$\mathrm{ReLU} = max\ (0, x) \tag{3}$$

Pool operation is performed after convolution is completed. Maximum pooling is helpful to extract the most significant features in the feature map, and it is invariant to small displacement and deformation [16]. This is because the maximum pool retains the maximum response in each pool area and ignores other smaller responses. Expressed by the following equation:

$$R_n = Max(Y_{ij}) \tag{4}$$

R_n represents the characteristic matrix of the sequence image after convolution and pooling operation on the nth image. After that, all the image frames in the sequence are operated in turn, and the final feature matrix can be expressed as $R = (R_1, R_2, R_3, \ldots, R_n)$, where $n \le N$ [17–19].

In this model, a three-layer CNN is used to process the spatial information of video frames. Each layer uses a convolution kernel with a size of 3 × 3, with a step of 1, and is filled with "SAME" to keep the feature map size unchanged. After convolution, the dimension of features is reduced by the maximum pool layer of 2 × 2, and the pool step is 0, which is helpful to extract salient features and enhance the robustness to small displacement and deformation. This structure ensures the effectiveness and computational efficiency of feature extraction.

3.2. Processing layer of LSTM model

An output *R* of CNN layer corresponds to a time point and serves as the input of LSTM, in which there is a memory unit *C* running through the whole sequence. This memory cell *C* is jointly regulated by the input gate, the forgetting gate, the output gate and the previous memory cell. For a certain moment t, the equation for updating the weight of each part of the LSTM unit is as follows:

Memory unit update:

$$\tilde{C}_t = tanh(W_{xc}X_t + U_{hc}h_{t-1} + b_c) \quad (5)$$

X_t represents the input characteristic matrix of time t. W_{xc} represents the weight matrix data from the input layer to the storage unit. U_{hc} represents the weight parameter matrix from the hidden layer to the memory cell. b_c represents the offset parameter values between memory cells. $tanh$ stands for hyperbolic tangent activation function [20].

Input gate update:

$$i_t = \sigma(W_{xi}X_t + U_{hi}h_{t-1} + W_{ci}C_{t-1} + b_i) \quad (6)$$

σ represents the nonlinear activation function of Sigmoid applied to this network. W_{xi} represents the weight matrix data from the input layer to the input gate. U_{hi} represents the weight parameter matrix from the hidden layer to the input gate. b_i represents the offset parameter value between the input gates [21].

Forget gate update:

$$f_t = \sigma(W_{xf}X_t + U_{hf}h_{t-1} + W_{cf}C_{t-1} + b_f) \quad (7)$$

W_{xf} represents the weight matrix data from the input layer to the forget gate. U_{hf} represents the weight parameter matrix from the hidden layer to the forget gate. b_f represents the offset parameter value between forget gates [22].

Output gate update:

$$o_t = \sigma(W_{xo}X_t + U_{ho}h_{t-1} + W_{co}C_t + b_o) \quad (8)$$

W_{xo} represents the weight matrix data from the input layer to the output gate. U_{ho} represents the weight parameter matrix from the hidden layer to the output gate. b_o represents the offset parameter value between the output gates [23].

Calculation of new memory cells:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (9)$$

$\odot$ represents a bitwise multiplication operation.

Hide status updates:

$$h_t = o_t \odot \tanh(C_t) \quad (10)$$

In LSTM, the memory cell C_t is calculated from the data of two parts: one part is the point multiplication result of the memory cell C_{t-1} and the forget gate f_t at the previous moment. The other part is the point multiplication result of the new memory cell $\tilde{C}_t$ calculated by the input gate i_t, the current input X_t and the state h_{t-1} at the previous moment. The activation value range of the input gate and the forgetting gate is (0–1). When the activation value is 0, it means that the previous information has been completely forgotten. When the activation value is not 0, it means that the current hidden layer will learn some information of the previous time unit [24]. Output gate o_t determines the ratio of memory information of the current cell to output information, and affects the contribution of long-term memory to output information. Because LSTM improves the processing ability of these gating mechanisms and cell information, and can learn long-term dependent information, the gradient will not disappear in a short time when training with back propagation algorithm, thus reducing the problem of gradient disappearance and making the model training smoother [25].

3.3. Prediction result output layer

To recognize video action data sets efficiently, a neural network structure with two fully connected layers is designed. First, the first fully connected layer receives all the time series data processed by the LSTM unit, which contains rich time series information. Through this step, the network can integrate the local characteristics of the action, thus capturing the dynamic change and continuity of the action. The weighted average of the features is then determined by statistically analyzing these mixed features once again in the second fully connected layer, where every node is fully connected to every other node in the top layer. This average value is a key statistic in the process of action recognition, which integrates information from different time points and provides a stable feature representation for the classifier. Finally, this average value is sent to the classifier and Softmax function to predict the probability of action categories, and its calculation process follows the equation:

$$y_t = W_t \times h_t \tag{11}$$

Through this structural design, the model can not only learn the time sequence characteristics of actions, but also realize the effective fusion and classification of features in the fully connected layer, thus improving the accuracy and robustness of action recognition.

4. Model performance evaluation

4.1. Experimental setup

4.1.1. Experimental environment

In this experiment, to ensure the efficiency and accuracy of table tennis players' technical action analysis, high-performance computing equipment is used in the experiment. The experimental equipment includes Melcore i7–8700 Central Processing Unit (CPU), the main frequency is 3.2 GHz, the internal memory is 16 GB, and GeForce GTX 1080 Ti Graphics Processing Unit (GPU). These configurations ensure high efficiency in processing a large number of video data and training complex models. Python 3.6 is chosen as the programming language, and PyTorch 1.4.1 is used as the deep learning framework. PyTorch's dynamic calculation diagram and convenient Application Programming Interface (API) call make the model construction and debugging more flexible and efficient. Related dependency libraries include Computer Unified Device Architecture (CUDA) 10.1, Computer Vision (OpenCV) 4.4 and Numerical Python (NumPy) 1.19.2, which play an important role in video data processing, numerical calculation and deep learning model training.

4.1.2. Dataset

Two main datasets are used in the experiment: Table Tennis Tracking Network (TTNet) dataset and self-created table tennis technical action video dataset. TTNet dataset is a public dataset widely used in action recognition. The self-created dataset focuses on the technical movements of table tennis players, including serving, forehand hitting, backhand hitting and spike, with a total of 565 videos, each of which is 2–3 s.

The combination of these datasets enables the model to be carefully trained in a specific sports scene, thus improving the recognition accuracy and feature extraction ability of table tennis technical actions. Adam optimizer is used in the model training process, and appropriate learning rate and batch size are set. To make sure the model can thoroughly understand and optimize the technical actions of table tennis, the number of training rounds is set at fifty. During the testing phase, the model's performance in real-world applications and its capacity for generalization are confirmed using a separate test set. Through the above experimental setup and data processing, this paper aims to deeply analyze the technical actions of table tennis players, improve the accuracy of action recognition by using deep learning technology, and provide effective technical analysis tools for table tennis players and coaches.

4.2. Model performance comparison

To comprehensively evaluate the proposed performance of the deep learning model in table tennis technical action analysis, this paper compares the performance of different models in technical action recognition tasks. The following are the results of comparing CNN, LSTM, Multi-Objective Function (MOF) and CNN-LSTM combined models. The contents of comparison include accuracy, precision, recall and F1 score. The comparison results of model performance are shown in **Table 1** and **Figure 2**.

Table 1. Model performance comparison.

Model	Accuracy	Precision	Recall	F1 Score
CNN	0.875	0.869	0.88	0.874
LSTM	0.892	0.887	0.895	0.891
MOF	0.802	0.795	0.808	0.801
CNN-LSTM	0.923	0.918	0.925	0.921

The performance comparison of different models reveals their significant differences in analyzing table tennis players' technical actions. The accuracy rate of CNN is 0.875, the precision rate is 0.869, the recall rate is 0.880 and the F1 score is 0.874. This shows that CNN is excellent in overall performance, but its index is slightly lower than that of more advanced models. The performance of LSTM is superior, with accuracy of 0.892, precision of 0.887, recall of 0.895 and F1 score of 0.891, showing a good balance between precision and recall. In contrast, the effect of MOF is poor, the accuracy rate is 0.802, the precision rate is 0.795, the recall rate is 0.808, and the F1 score is 0.801, which shows that MOF is not as applicable as other models in this task. It is worth noting that the CNN-LSTM hybrid model performs the best, with an accuracy of 0.923, a precision of 0.918, a recall of 0.925 and a F1 score of 0.921, which is excellent in all indicators. This highlights that CNN-LSTM model can effectively combine the advantages of convolution and sequence learning, and provides the most accurate and balanced results for table tennis technical action analysis.

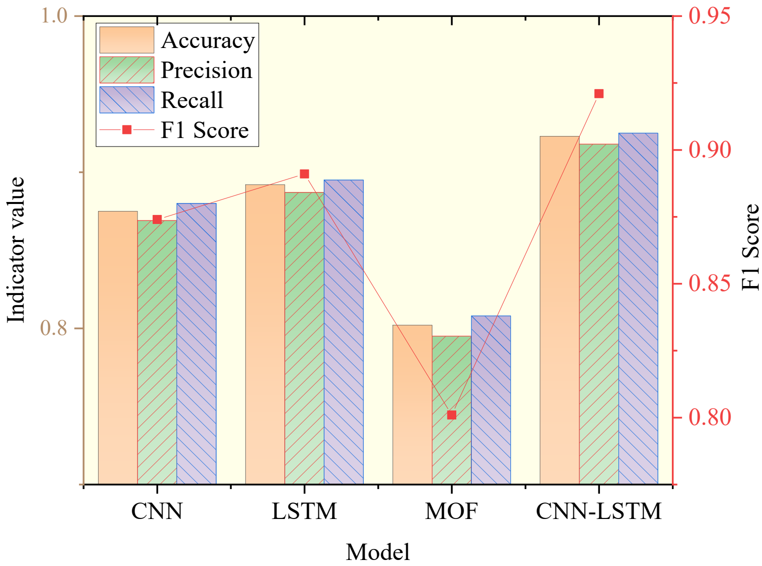

Figure 2. Model performance comparison results.

4.3. Technical action analysis

The paper identifies common table tennis player technical actions (e.g., serving, forehand, backhand, spike, etc.) to assess the model's effectiveness in classifying various technical actions. In table tennis competition, the time sequence characteristics of technical actions are very important for understanding athletes' competition strategies and technical application. The sequence consistency and classification accuracy of actions are shown in **Table 2** and **Figure 3**.

Table 2. Timing consistency and classification accuracy of actions.

Technical Action	Classification Accuracy	Temporal Consistency
Serve	0.930	0.895
Forehand Stroke	0.915	0.872
Backhand Stroke	0.92	0.884
Smash	0.907	0.867

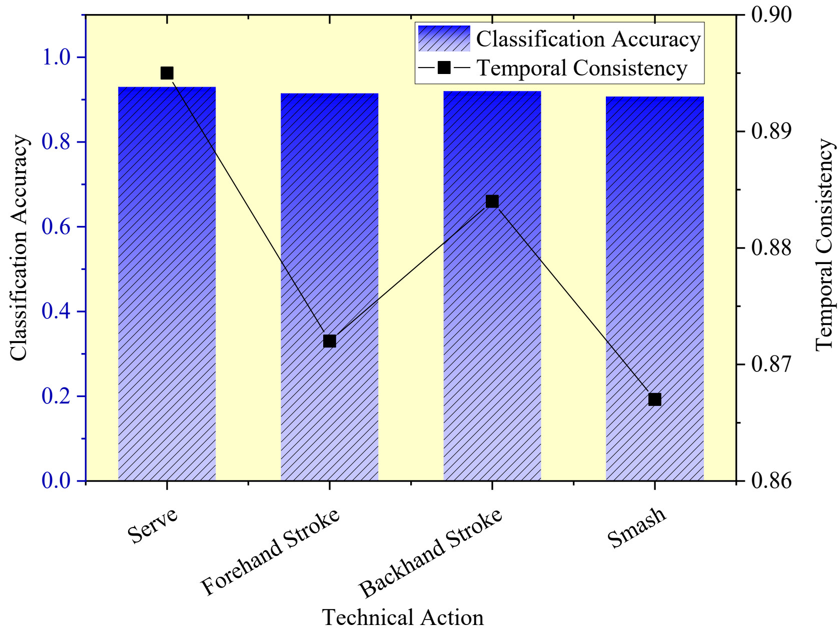

Figure 3. Timing consistency and classification accuracy of actions.

According to the data in the figure, the model has the highest classification accuracy at the level of service technical action, reaching 0.930, which shows that the recognition ability of service action is particularly excellent. This high accuracy not only reflects the superior performance of the model in service action recognition, but also reveals the unique advantages of service technology itself in characterization. Serve action usually has strong characteristics, including starting posture, hitting point and swing trajectory, which are important basis for effective recognition of the model. Specifically, the starting posture of serving is usually standardized, and the athlete's preparation before serving can provide clear visual information for the model. In addition, the height and direction of the hitting point have significant characteristics in different types of serve, which help the model to quickly identify different serve strategies and techniques. The trajectory of the swing is the dynamic performance of the serve. By analyzing the speed, angle and path of the swing, the model can better understand the technical details of the athletes in serving. Therefore, the model can capture these key features well and make it perform well in serve recognition. This ability to accurately identify the serve action not only enhances the application potential of the model in technical analysis, but also provides more accurate training guidance and tactical adjustment suggestions for coaches and athletes.

The accuracy of forehand stroke and backhand stroke is 0.915 and 0.920 respectively, which also shows that the model performs well in the classification of these two common strokes. The accuracy of spiking is slightly lower, which is 0.907, but it remains at a high level. In terms of time sequence consistency index, the time sequence consistency of service is 0.895, which also shows high stability. The consistency of forehand stroke and backhand stroke is 0.872 and 0.884, respectively, which shows that the model can capture the timing characteristics of these actions well. The consistency of spike sequence is the lowest, which is 0.867, but it still maintains a high consistency. These results show that the model can not only accurately classify table tennis players' technical actions, but also effectively analyze the time sequence characteristics of technical actions, thus providing valuable data support for athletes' technical analysis and tactical adjustment.

4.4. Real-time feedback performance evaluation

The paper evaluates the deep learning model's real-time analysis capability to assess the system's real-time feedback performance in a real-world setting. In order to verify the model's performance in a real-time setting, this paper captures the processing time and feedback delay using video footage from various tournaments. The real-time feedback performance is shown in **Figure 4**.

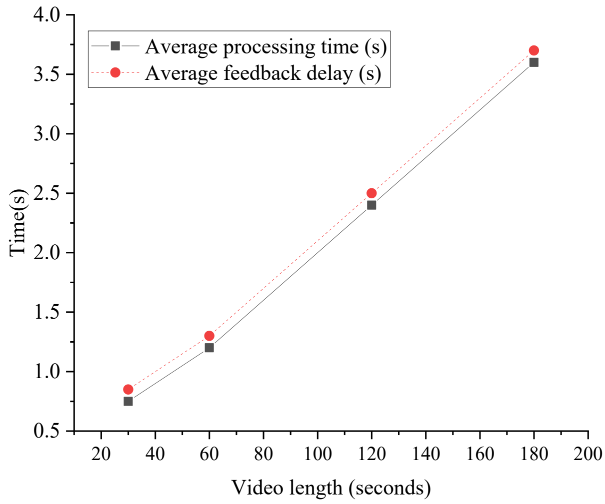

Figure 4. Real-time feedback performance.

The findings indicate that both the average processing time and average feedback delay exhibit a linear growth pattern as the length of the video increases. When the video length is 30 s, the average processing time is 0.75 s and the average feedback delay is 0.85 s. When the video length increases to 60 s, the processing time and feedback delay increase to 1.20 s and 1.30 s respectively. For a 120-s video, the processing time is 2.40 s, and the feedback delay is 2.50 s, while for a 180-s video, the processing time is 3.60 s, and the feedback delay is 3.70 s. These data show that the overall performance of the model still meets the requirements of real-time analysis, which provides reliable support for timely feedback in practical applications, although the time and delay of processing longer videos have increased.

5. Conclusion

The paper proposes a deep learning-based table tennis technical action analysis model, and several experiments confirm its superiority and efficacy. Following a comparative and analytical analysis of the CNN, LSTM, MOF, and CNN-LSTM combined models, the CNN-LSTM mixed model emerges as the most effective model across all metrics, with particular emphasis on accuracy, precision, recall, and F1 score. This model's outstanding performance demonstrates the potent benefits of combining CNN and LSTM in sequence data processing. Furthermore, the classification and timing consistency of several table tennis technical actions are also thoroughly analyzed in this paper. The experimental results demonstrate that while detecting serve, forehand stroke, backhand stroke, and spike, the model can retain high classification accuracy and consistency of time sequence. This offers solid data support for the tactical and technical analysis of table tennis players. Ultimately, by assessing the system's real-time feedback performance, it is determined that the model possesses a robust real-time analysis capability and can deliver efficient and timely feedback in real-world scenarios. Although this paper has made remarkable achievements, there are still some limitations. Firstly, the model training process consumes a lot of computing resources and requires high hardware, which may limit its application in the resource-limited environment. Secondly, the diversity and quality of datasets need to be further improved to enhance the generalization ability and accuracy of the model. Future research can further optimize the computational efficiency of the model to reduce the demand for hardware resources, and explore its application in other sports to verify its universality in different sports scenes. In addition, improving the diversity and quality of datasets will also be an important direction to enhance the generalization ability of the model and provide more extensive support for intelligent sports analysis.

Ethical approval: Not applicable.

Conflict of interest: The author declares no conflict of interest.

References

1. Sun H. Improving the Accuracy of Recognition and Evaluation of Technical Movements of Basketball Players Using Deep Learning Algorithms. Journal of Electrical Systems. 2024; 20(6s): 1959–1969.
2. Rong Z. Optimization of table tennis target detection algorithm guided by multi-scale feature fusion of deep learning. Scientific Reports. 2024; 14(1): 1401.
3. He S. The Application of Deep Learning-Based Technique Detection Model in Table Tennis Teaching and Learning. Systems and Soft Computing. 2024; 1(1): 200116.
4. Lyu L, Huang Y. Sports activity (SA) recognition based on error correcting output codes (ECOC) and convolutional neural network (CNN). Heliyon. 2024; 10(6):23–43.
5. Chen ZH, Zhang GQ. CNN sensor based motion capture system application in basketball training and injury prevention. Preventive Medicine. 2023; 174(12): 107644.
6. Gao Y, Qian Q, Sun X, et al. Evaluation model of athletes' lower extremity training ability based on LSTM algorithm. Int. Arab J. Inf. Technol. 2024; 21(1): 147–157.
7. Khobdeh SB, Yamaghani MR, Sareshkeh SK. Basketball action recognition based on the combination of YOLO and a deep fuzzy LSTM network. The Journal of Supercomputing. 2024; 80(3): 3528–3553.
8. Song H, Li Y, Fu C, et al. Using complex networks and multiple artificial intelligence algorithms for table tennis match action recognition and technical-tactical analysis. Chaos, Solitons & Fractals. 2024; 178(2): 114343.
9. Yenduri S, Chalavadi V. Adaptive temporal aggregation for table tennis shot recognition. Neurocomputing. 2024; 584(1): 127567.
10. Zhang J, Ren Y, Lin L, et al. Table tennis motion recognition based on the bat trajectory using varying-length-input convolution neural networks. Scientific Reports. 2024; 14(1): 3549.

11. Host K, Pobar M, Ivasic-Kos M. Analysis of movement and activities of handball players using deep neural networks. Journal of imaging. 2023; 9(4): 80.
12. Zhao L. A hybrid deep learning-based intelligent system for sports action recognition via visual knowledge discovery. IEEE Access. 2023; 11(8): 46541–46549.
13. Cheng L. Athletes' Action Recognition and Posture Estimation Algorithm Based on Image Processing Technology. Journal of Electrical Systems. 2024; 20(6s): 2059–2069.
14. Li X, Ullah R. An image classification algorithm for football players' activities using deep neural network. Soft Computing. 2023; 27(24): 19317–19337.
15. Liu L. Objects detection toward complicated high remote basketball sports by leveraging deep CNN architecture. Future Generation Computer Systems. 2021; 119(1): 31–36.
16. Li W. Application of IoT-enabled computing technology for designing sports technical action characteristic model. Soft Computing. 2023; 27(17): 12807–12824.
17. Omarov B, Nurmash N, Doskarayev B, et al. A Novel Deep Neural Network to Analyze and Monitoring the Physical Training Relation to Sports Activities. International Journal of Advanced Computer Science and Applications. 2023; 14(9): 22–31.
18. Leddy C, Bolger R, Byrne PJ, et al. The application of Machine and Deep Learning for technique and skill analysis in swing and team sport-specific movement: A systematic review. International Journal of Computer Science in Sport. 2024; 23(1): 110–145.
19. Sarwar MA, Lin YC, Daraghmi YA, et al. Skeleton Based Keyframe Detection Framework for Sports Action Analysis: Badminton Smash Case. IEEE Access. 2023; 11(12): 90891–90900.
20. Wang TY, Cui J, Fan Y. A wearable-based sports health monitoring system using CNN and LSTM with self-attentions. Plos one. 2023; 18(10): e0292012.
21. Sun X, Wang Y, Khan J. Hybrid LSTM and GAN model for action recognition and prediction of lawn tennis sport activities. Soft Computing. 2023; 27(23): 18093–18112.
22. Li X, Luo R, Islam FU. Tracking and detection of basketball movements using multi-feature data fusion and hybrid YOLO-T2LSTM network. Soft Computing. 2024; 28(2): 1653–1667.
23. Pajak G, Krutz P, Patalas-Maliszewska J, et al. An approach to sport activities recognition based on an inertial sensor and deep learning. Sensors and Actuators A: Physical. 2022; 345(2): 113773.
24. Zhang Q, Zhang X, Hu H, et al. Sports match prediction model for training and exercise using attention-based LSTM network. Digital Communications and Networks. 2022; 8(4): 508–515.
25. Petrov E, Ivanov D. Neural Signal Denoising in Sports Biomechanics: Dynamic Arrow Characterization with Stacking Ensemble of Hybrid CNN-RNN. Journal of Humanities and Applied Science Research. 2022; 5(1): 89–103.