ISTP

RESEARCH ARTICLE

Motion Capture Algorithm for Students' Physical Activity Recognition in Physical Education Curriculum

Yinfu Lu[1,2] and Haitao Long[1,2]

[1]National University (Philippines), Manila, Philippines
[2]Physical Education Institute, Hunan Institute of Science and Technology, China

Abstract: Physical training learning is one of the important ways to raise the national physical quality and health level. However, there are many problems in traditional physical education, such as the difficulty to identifying the effectiveness of physical education curriculum and low level of repetitive exercise content. In order to solve this problem and improve the curriculum quality of current middle school physical education courses, a motion capture algorithm (MCA) based on convolutional neural network and long-term memory network is proposed, and a student physical activity capture model is constructed based on the fusion algorithm. In the performance comparison test of the fusion algorithm proposed in this study, the loss value and accuracy of this fusion algorithm are 0.045 and 0.921, respectively, significantly superior to the comparison algorithm. Then in the empirical analysis, the accuracy rate of this MCA model proposed in this study for students' walking posture recognition in physical education courses is 91.5%, which is better than the comparative capture method. This MCA can accurately capture the physical activities of students in physical education courses, which has practical application significance.

Keywords: CNN; gesture recognition; LSTM; motion capture; physical education curriculum

I. INTRODUCTION

As society develops and people's awareness of health increases, physical education curriculum (PEC) plays an important role in school education [1]. PEC not only helps students' physical health development but also cultivates their teamwork spirit and confidence. However, traditional physical education teaching models often rely on teachers' subjective judgments and students' self-awareness, which have a certain degree of subjectivity and inaccuracy [2]. To better guide students' physical training and improve teaching effectiveness, it is of great significance to find an excellent motion capture algorithm (MCA) [3,4]. Convolutional neural network (CNN) is a deep learning algorithm commonly used in artificial neural networks to process data with grid structures [5]. CNN is successfully used in image recognition, object detection, and speech recognition [6,7]. Compared with traditional machine learning methods, CNN can automatically learn more abstract and advanced feature representation from the original data, thus improving the accuracy and robustness of classification and recognition [8–10]. To better recognize the physical interaction of students in PEC, this study combines CNN and long short-term memory (LSTM) and applies them to an action recognition model. It is expected that this model can accurately identify and analyze students' physical activities, thereby monitoring students' movement status, posture, and action effects in real time and improving PEC effectiveness. The purpose of this study is to develop an MCA for physical activity of students in physical education courses. Compared with traditional MCAs, this algorithm has higher accuracy and precision and can identify the physical activity of students more accurately. In addition, this research has innovation and

contributions in the aspects of algorithm innovation, educational application, student personalized guidance, and data analysis research, etc., which is expected to bring important progress and improvement to the field of physical education and sports training. This paper is divided into the following sections. Section II discusses the related works. Section III.A analyzes the MCA and the PEC evaluation model. Section III.B describes the process of constructing the MCA model of PE students with CNN-LSTM algorithm. Section IV compares and analyzes the performance of CNN-LSTM MCA and PE students' motion capture model. Section V summarizes the full text.

II. RELATED WORKS

Physical education is not only an important part of quality education but also an important basis for comprehensive education. With the development of AI technology, MCAs have been extensively studied in various fields. To better identify human movement, Kanko *et al.* applied a new spatio-temporal parameter and compared the measured results by Bland–Altman method, mean difference, Pearson correlation coefficient, and intra-group correlation coefficient. Through experimental detection, the unmarked motion capture can fully measure the spatiotemporal gait parameters of healthy young people during exercise [11]. In order to improve the accuracy of human motion recognition in small- and medium-sized video ranges and the computational efficiency of large-scale datasets, Gao *et al.* established multidimensional data models of motion recognition and motion capture based on video images with deep learning framework. After large-scale simulation data experiments, the results showed that the accuracy rate of human behavior classification was 89.79% [12]. To avoid ghost and parallax effects in image recognition, Laaroussi *et al.* replaced noise function with

Corresponding author: Haitao Long (e-mail: longhaitao1991@163.com).

1

fractional Brownian motion of predetermined similarity function to detect dynamic object actions in public areas. After experimental analysis, this method could quickly and effectively find the best seam and avoid the occurrence of pseudo-ghost [13]. In order to evaluate the accuracy and processing time of CT-3DUS registration algorithm, Chan's team adopted the motion capture 3D ultrasound (3DUS) system to 3D print and scan human vertebrae to guide posterior spinal surgery. After 191 tests of image registration, the system achieved the accuracy threshold required for surgery [14]. To effectively prevent chronic pain and disability, Onks *et al.* used micro-Doppler signals generated by micro-Doppler radar to identify movement differences in military environments. Experimental results showed that compared with human eyes, micro-Doppler signals could effectively identify more subtle movement patterns [15].

In physical exercise, personal exercise can not only effectively strengthen the physique but also provide a material basis for a healthy personality. The evaluation of students' physical behavior and movement can effectively evaluate the quality results of physical education courses. To effectively classify physical activities, Manuel *et al.* divided the acceleration signals into overlapping windows and detected the execution activities of each window based on CNN. After experimental detection, the significant performance of human activity recognition increased from 89.83% to 96.62% [16]. Wanting *et al.* used multiple logistic regression to analyze the correlation between physical activity and physical achievement in PE class in order to examine the relationship between physical activity and physical achievement of students. From the sample testing, the physical activity status of students significantly affected physical achievement [17]. To improve the accuracy of students' classroom behavior recognition algorithm, Chonggao *et al.* adopted cluster analysis algorithm and random forest algorithm to improve the traditional algorithm and build a network topology model. After data detection, they found that the proposed network topology model was superior to the network structure of a single feature extraction algorithm [18]. In order to identify normal activity and brute force attack activity, Randhawa's team used decision tree, k-NN classifier, support vector machine, and wearable inertial fabric sensor to conduct various experiments on machine learning (supervised) classification technology. The experimental results showed that the vector algorithm provided 97.6% accuracy and 0.85 seconds of computation time for activity classification [19]. To give feedback on the performance of body movements in physical activities, Ferreira *et al.* proposed a validation system based on 2D human pose estimation networks and deep semantic features. Through fitting exercises, the model was correct more than 92% in 4 out of 5 exercises [20].

To sum up, the motion capture technology has become an important research topic in the field of computer vision and has been widely applied in behavior recognition and video surveillance by experts and scholars. The recognition of students' movements in physical education is related to students' participation in class and sensitivity of body function and affects teachers' recognition and training of students' physical quality. At present, there are few studies on motion capture technology for motion recognition of students in physical education. Therefore, this paper attempts to construct a motion capture system based on CNN-LSTM algorithm and carries out related experimental research.

III. MCA FOR STUDENT PHYSICAL ACTIVITY RECOGNITION IN PEC

In the PEC, the accurate identification of students' physical activity is important for teachers and students' physical education and training. However, the traditional MCA has some limitations in sports scenes, which cannot meet the needs of accurate recognition and real-time monitoring of students' movements. Therefore, this study aims to develop an MCA for physical activity recognition of students in physical education courses. The algorithm will realize the high-precision recognition and real-time monitoring of students' movements by using CNN algorithm and LSTM algorithm. Through personalized instruction and feedback, the algorithm will help students improve movement skills and improve sports performance, providing strong support for the development of physical education and sports training.

A. MCA CONSTRUCTION BASED ON CNN-LSTM ALGORITHM FUSION

CNN is a deep learning model, which is successfully used in different fields and has become one of the most important models in deep learning [21,22]. The advantage of CNN lies in its ability to automatically learn features in images, as well as its translation invariance and local connectivity, making it excellent in processing images and other two-dimensional data [23,24]. Figure 1 shows the structure and parameters of CNN.

In Fig. 1, the input image is added to the CNN input layer. After each convolution layer, the features are extracted to obtain the feature images C1, C2, C3, C4, and C5. After multiple convolution and pooling, these images are classified or returned through the fully connected layer. The fully connected layer means that every input neuron is connected to an output neuron. In this layer, each feature of the input data is computed with all the output neurons.
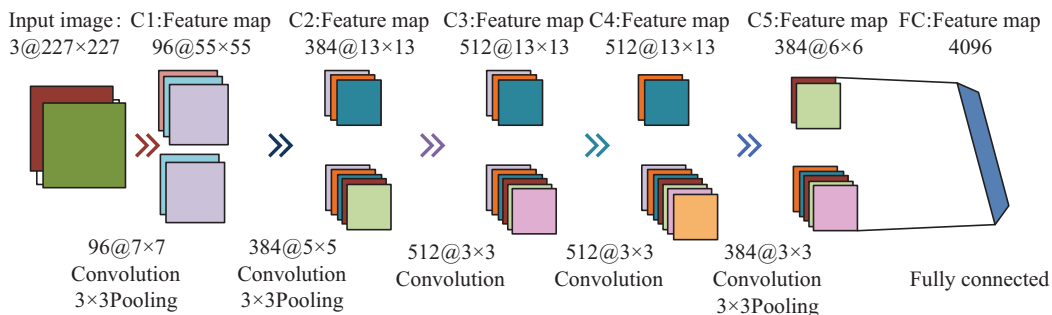

Fig. 1. Structure and parameters of CNN.

It is usually used to connect the feature extraction results of the previous layer with the output layer to make the final prediction. CNN can gradually extract higher-level features by stacking multiple convolutional and pooling layers, and equation (1) is the calculation of its convolutional layers.

$$x^l_j = f\left(\Sigma^n_{i \in M_j} x^{l-1}_i * k^l_{i,j} + b^i_j\right) \qquad (1)$$

In equation (1), x^l and f are the output and activation function of lth layer, respectively. x^{l-1}_i is the output of layer $l-1$. k^l and b^l_j are the lth convolution kernels and offset terms. M_j is the selected input feature set. Equation (2) is the calculation of pooling layer.

$$a = \begin{bmatrix} 2 & 3 & 0 & 3 \\ 1 & 4 & 4 & 3 \\ 5 & 6 & 4 & 3 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad a_{max} = \begin{bmatrix} 4 & 4 \\ 6 & 4 \end{bmatrix},$$

$$a_{ave} = \begin{bmatrix} 2.5 & 2.25 \\ 3 & 2 \end{bmatrix} \qquad (2)$$

Finally, these features are mapped to different categories or regression targets through a fully connected layer. In addition, the research improves it based on activation function, regularization method, random gradient descent method, and batch normalization method to increase MCA reliability. Among them, RELU is chosen as the activation function, which can reduce the linear results of CNN output. Compared with other activation functions, this function can also reduce the frequency of saturation and increase convergence rate. The regularization method used in the study is dropout, which can effectively reduce over-fitting problems and increase the robustness of CNN [25,26] in equation (3).

$$H(p,q) = -\sum_x p(x) \log q(x) \qquad (3)$$

In equation (3), $p(x)$ stands for the target distribution. $q(x)$ refers to the predicted distribution. The random gradient descent can optimize neural network training. It takes the gradient as the expectation, and samples the total data with small samples, thus reducing the total amount of extracted data and the gradient update time [27,28]. The batch normalization method is applied to convolution layer to reduce features correlation obtained from training and then accelerate its convergence rate. The motion capture object proposed in this study is the video data of students' physical activities. To conduct time series related recognition research on videos, this study combines CNN and LSTM to construct an MCA based on CNN-LSTM. LSTM is a special recurrent neural network,

which can be used to deal with gradient disappearance and gradient explosion encountered by traditional RNN when processing long sequences [29]. LSTM can realize long-term memory and control information flow by introducing memory unit and gating mechanism [30]. Equation (4) is the calculation of forgetting gates.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \qquad (4)$$

h_{t-1} represents the previous hidden layer's output in equation (4). σ is a sigmoid function. x_t represents the current input. W and b represent the weight matrix and bias. Figure 2 shows the proposed CNN-LSTM MCA structure.

In Fig. 2, the CNN-LSTM MCA proposed in this study includes a CNN-based spatial flow network module and an LSTM-based temporal flow module. This study trained the spatial flow module and temporal flow module separately. In the spatial flow module, research will extract RBG data from action data as input to the spatial flow network for spatial training. After inputting RBG data into the spatial flow module, the data are feature extracted using CNN and iteratively calculated based on the sum of time series features using LSTM. The operation process of time flow module is similar to the space flow module. However, its input is the optical flow field form data extracted from the action data. Finally, this study weighted and fused the spatial flow network module and the time flow network module through category scores to achieve the recognition and capture of student actions.

B. INTEGRATING IMPROVED MCA INTO A PEC STUDENT BODY ACTION RECOGNITION MODEL

The study of student body movement recognition in PEC refers to using computer vision and machine learning techniques to automatically recognize and analyze the body movements performed by students on PEC [31,32]. The development of this research has made certain progress and has broad application prospects in education. In this study, CNN-LSTM MCA is constructed and applied to PEC, and a PEC student action recognition model based on fusion algorithm is constructed. It is expected to improve the effectiveness and accuracy of students' body movement recognition. Figure 3 shows the overall framework of this model.

In Fig. 3, this study inputs student action recognition data into a student action recognition model, which extracts data features in both optical flow and RBG frame formats, and divides the data into training and testing sets. Subsequently, the training set data are preprocessed. The optical flow dataset RBG data from the training set are input into spatial and time flow network modules for training. Further spatial flow network optimization during the model construction process is studied, and residual networks,
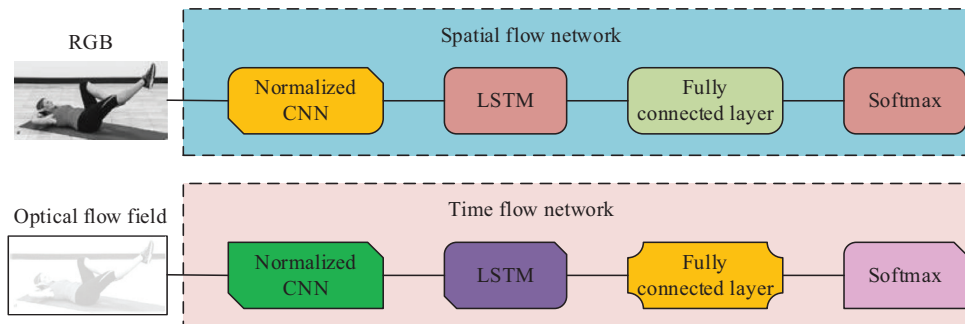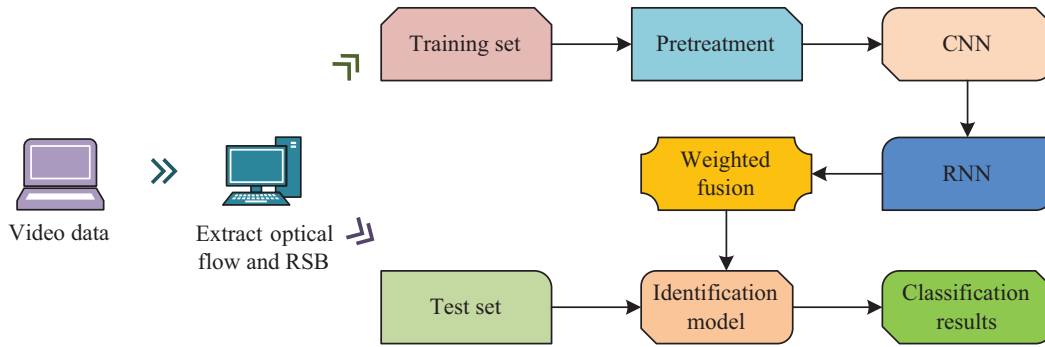

Fig. 2. CNN-LSTM MCA.

Fig. 3. Construction process of student action recognition model in physical education classroom.

concat layers, and local corresponding normalization layers are introduced into the spatial flow network. Among them, the introduction of residual network is used to solve the problem of vanishing gradient problem of spatial flow network, and the introduction of concat layer is to reduce over fitting. Finally, the local corresponding normalization layer can distinguish them based on the strength of neuron feedback response, thus enhancing the generalization of spatial flow network. Equation (5) is a normalized calculation.

$$x^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

In equation (5), x^* represents the normalized data. x_i is the ith data in dataset. $x_{\min}$ and $x_{\max}$ are the minimum and maximum values in dataset, respectively. Finally, these two modules are weighted and fused to construct an action recognition model. Finally, the test set data are input into the recognition model and the action recognition results are obtained. Figure 4 shows a weighted fused student action recognition model.

In Fig. 4, the optical flow data include horizontal and vertical optical flows in a plane coordinate system. In the study, the optical flow data extraction method is used to reduce the scale of horizontal and vertical optical flow data to [128, +128]. Subsequently, this study combines vertical and horizontal optical flow data into three channels of optical flow and combines them with RGB frame data extracted from action video data. The extracted long video images are preprocessed and randomly edited into multiple frame sequences. Then, these two edited data streams are inputted into their respective CNNs through three channels, and the obtained

vector features are input into LSTM at both levels. Then, feature fusion is performed on the output data of CNN and LSTM through full connectivity, and a softmax classifier is used to perform probability averaging on the predicted labels in each frame to obtain video segment classification results for RGB frames and optical flow data. Finally, this study weights and merges two classification results obtained by training two data streams separately in a certain proportion to achieve human motion recognition.

IV. PERFORMANCE COMPARISON AND EMPIRICAL ANALYSIS OF CNN-LSTM ALGORITHM

To analyze the practical application effect of the CNN-LSTM algorithm proposed in the study and the student action recognition model based on the CNN-LSTM algorithm, the superiority of the CNN-LSTM algorithm is demonstrated through comparative analysis in this chapter, and the diagnostic accuracy of the fault diagnosis model is verified.

A. PERFORMANCE COMPARISON AND ANALYSIS OF CNN-LSTM MCA

To better recognize students' physical movements in PEC, CNN, and LSTM algorithms are fused to obtain CNN-LSTM MCA. To analyze the actual performance of CNN-LSTM MCA, the research compares it with LSTM algorithm, RCNN and CNN, and takes loss curve, accuracy rate, receiver operating characteristic (ROC), and
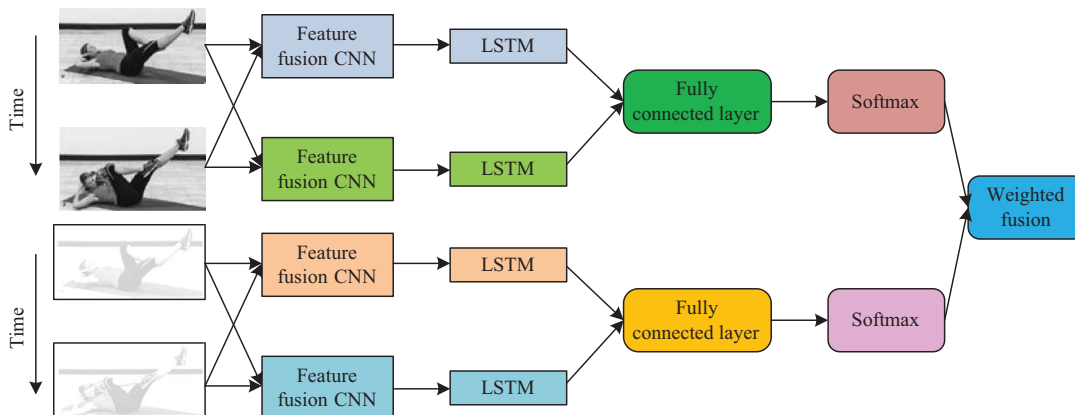

Fig. 4. PEC student action recognition model.

accuracy rate as comparison indicators. The dataset used in this comparison experiment is KTH dataset, which is a classic motion recognition dataset and is often used in the research of MCAs. It contains 2,391 sets of data covering six different actions, each performed by 25 people in four different scenarios. Thus, the KTH dataset has a total of 600 video sequences, and each can be further segmented into 4 subsequences. The motion of this dataset is standardized and the fixed lens is used, which is relatively rich for the motion recognition task. The performance of these four algorithms is compared by comparing each index on the KTH dataset. The experimental environment for this comparative experiment is as follows. The graphics card model used in this study is NVIDIA Ge Force GTX 1070, with a graphics memory capacity of 8GB. The computer CPU model is Intel i3-7100, with a main frequency of 3.9GHz and a memory of 8GB. Windows 7 is used as the computer system, and simulation is conducted using Matlab2017 software. Figure 5 shows the loss curve results and accuracy results of four algorithms.

In Fig. 5(a), as iteration increases, the maximum loss values of all four algorithms show a decreasing trend, and the minimum loss value of CNN-LSTM algorithm is 0.045. It is lower than LSTM algorithm's 0.088, RCNN's 0.124, and CNN's 0.193. In Fig. 5(b), as iteration increases, the accuracy of all four algorithms shows an upward trend, and CNN-LSTM algorithm has the highest accuracy after stabilization, at 0.921. It is higher than the LSTM algorithm's 0.863, RCNN's 0.815, and CNN's 0.796. In the dimensions of loss curve and accuracy, the combined CNN-LSTM algorithm's performance is superior to that of other comparison algorithms. Figure 6 shows the accuracy and ROC of four algorithms.

In Fig. 6(a), among four algorithms, the accuracy of CNN-LSTM algorithm is significantly better than other three comparative algorithms. The average accuracy of CNN LSTM algorithm is 94.3%. It is 89.6% higher than LSTM algorithm, 80.3% higher than CRNN algorithm, and 75.8% higher than CNN algorithm. In Fig. 6(b), among four algorithms, the ROC area of CNN-LSTM algorithm is larger than other three comparative methods. The ROC area of CNN-LSTM algorithm is 0.82, which is higher than 0.76 of LSTM algorithm, 0.64 of CRNN algorithm, and 0.61 of CNN algorithm. In terms of accuracy and ROC dimensions, the combined CNN-LSTM algorithm's performance is superior to that of
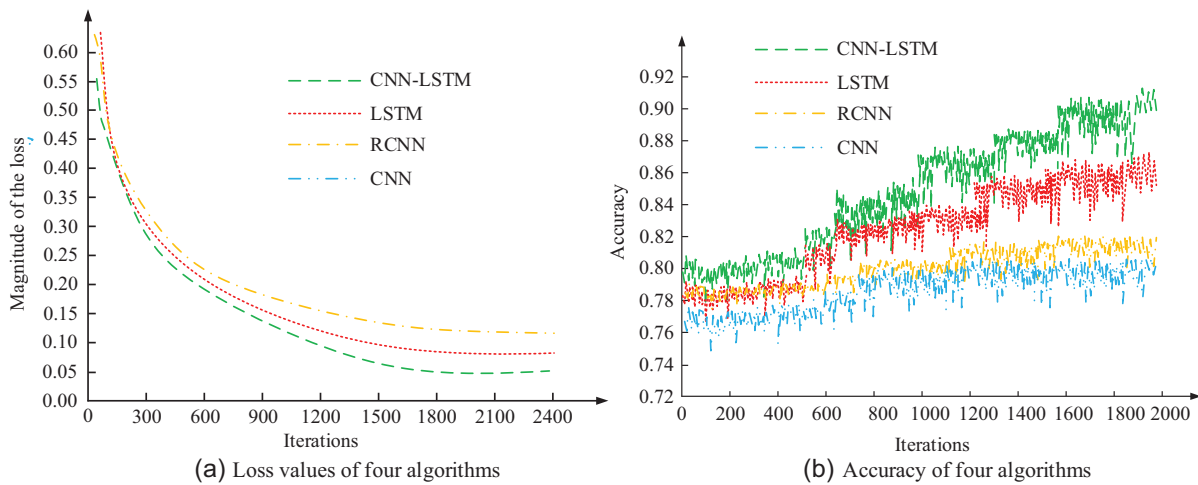

(a) Loss values of four algorithms

(b) Accuracy of four algorithms

Fig. 5. Loss curve results and accuracy results of four algorithms.

(a) Precision of the six experiments

(b) Receiver operating characteristic of four algorithms
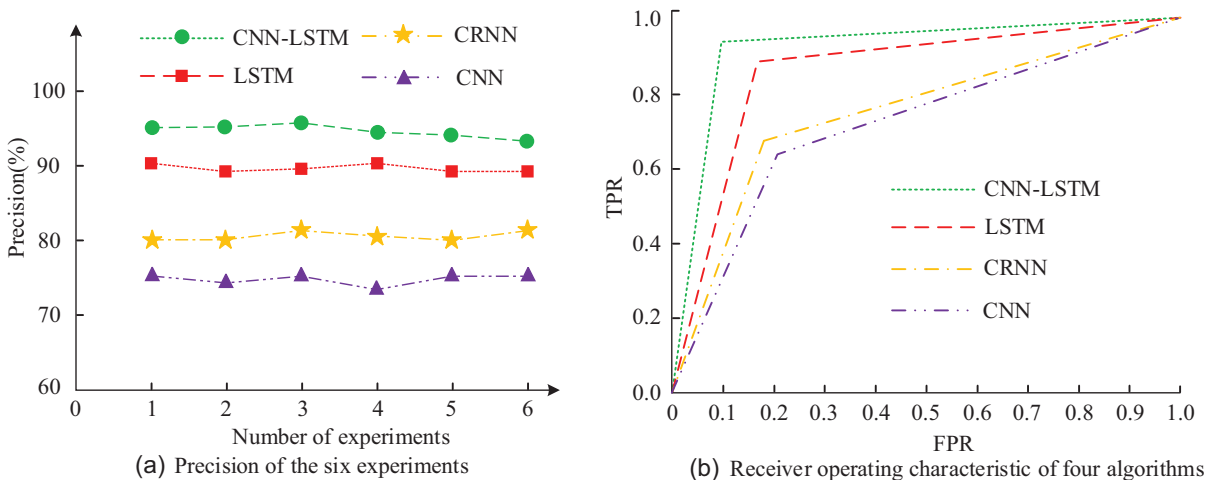
Fig. 6. Accuracy and ROC results of four algorithms.

other comparison algorithms. Then, the Precision-Recall (PR) curve and Area Under Curve (AUC) area of the four algorithms are carried out. The results of the PR curve and AUC curve of the four algorithms are shown in Fig. 7.

Figure 7(a) shows the PR curve of the four algorithms, and the PR curve represents the relationship between the accuracy rate and the recall rate. From Fig. 7(a), CNN-LSTM completely covers the other three algorithms, with the maximum area of 0.91. It can be seen that the learning performance of CNN-LSTM algorithm is better than the other three algorithms, and the area of CNN algorithm, Region-based Convolutional Neural Networks (RCNN) algorithm, and Long Short-Term Memory Network (LSTM) algorithm are 0.52, 0.62, and 0.76 in order. Figure 7(b) shows the AUC curve. AUC is an evaluation index to measure the quality of the binary classification model, and the probability that the positive cases are in front of the negative cases is predicted. From Fig. 7(b), the area of CNN-LSTM is 0.82, which is the maximum value. The area of CNN algorithm is 0.49, the area of RCNN algorithm is 0.9, and the area of LSTM algorithm is 0.62. The CNN-LSTM algorithm has better classification accuracy and learning ability, followed by LSTM algorithm and CNN algorithm.

Based on the comparison of these above dimensions, the combined CNN-LSTM algorithm's performance is far superior to LSTM algorithm, RCNN, and CNN. Therefore, applying it to PEC student action recognition can improve the accuracy of action recognition.

B. PERFORMANCE TESTING OF IMPROVED MCA

The above content provides a comparative analysis of the performance of CNN-LSTM MCA proposed in this study. To further analyze the practical application effect of this MCA, five classes of students from a certain junior high school are selected as the research subjects. In the experiment, five classes of students are selected as recognition objects for their walking, jogging, running, boxing, waving, clapping, and other postures in PEC. CNN-LSTM MCA, the traditional motion capture method, and image-based motion capture method are used to recognize students' motion posture in PEC, respectively. Figure 8 shows the accuracy results of three different methods for students' motion capture.

In Fig. 8, among these three methods, CNN-LSTM MCA has a higher recognition accuracy for students' various motion postures in PEC than other two motion recognition algorithms. And its
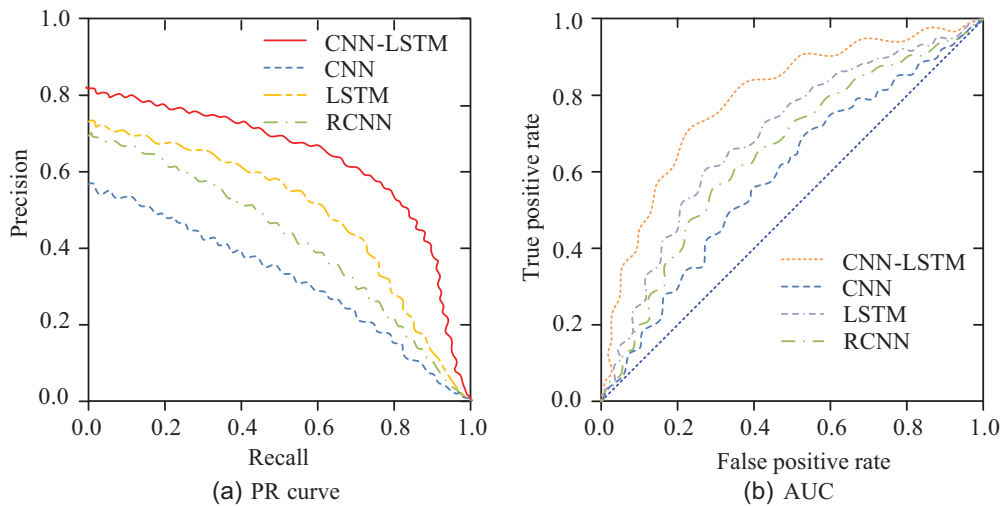

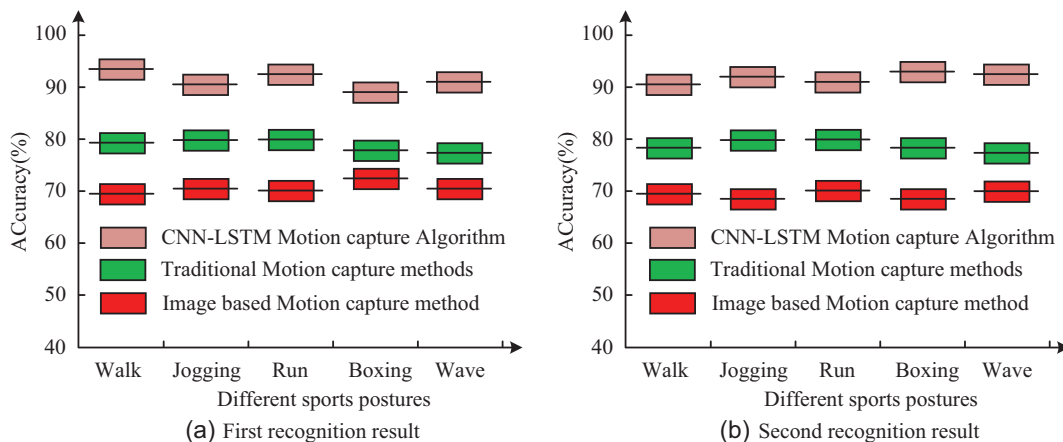**Fig. 7.** AUC and PR curves of four algorithms.

Fig. 8. Accuracy results of three different methods for students' motion capture.

recognition accuracy for walking posture is 91.5%. The recognition accuracy for jogging posture is 89.8%. The recognition accuracy for running posture is 90.4%. The recognition accuracy of boxing posture is 88.8%. The recognition accuracy for waving posture is 90.1%. The CNN-LSTM MCA proposed in this study has high recognition accuracy for various postures of PEC students and is superior to the comparative action recognition method. In addition, the research uses the confusion matrix to represent the proposed CNN-LSTM MCA's recognition results in the above dataset in Fig. 9.
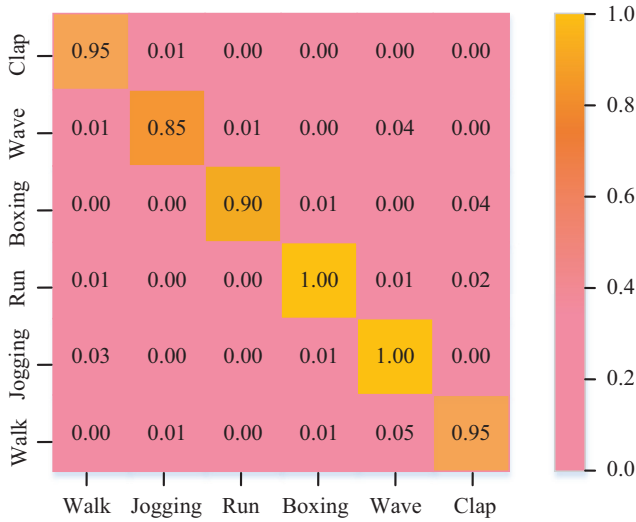

Fig. 9. Confusion matrix of action recognition based on student attitude dataset.

From Fig. 9, the algorithm has a low confusion rate for students' gestures. The CNN-LSTM MCA has good performance in terms of the dataset selected in the study. Therefore, the practical application of CNN-LSTM MCA is good, and its application in the movement recognition of students in physical education courses can effectively identify students' movements and promote the good development of physical education courses. Then in order to test the superiority of CNN-LSTM model, we start from four indexes: AUC curve, precision rate, recall rate, and accuracy rate. CNN-LSTM is compared with single algorithm of CNN and LSTM and mixed algorithm of Video-STM and DC-LSTM. The result is shown in Fig. 10.

The bar chart in the left half of Fig. 10 shows the comparative analysis between CNN-LSTM and other single algorithm models. As can be seen from the figure. Compared with single CNN algorithm and LSTM algorithm, DC-LSTM is superior in terms of accuracy, accuracy, AUC, and recall rate. Among them, in terms of recall rate and accuracy, DC-LSTM is outstanding, with a recall rate of 0.94 and an accuracy rate of 0.92. The right half of Fig. 10 shows the performance analysis of DC-LSTM compared with other hybrid models. It can be intuitively seen from the figure that the DC-LSTM algorithm exceeds the other two hybrid models in the four indexes of AUC curve, precision rate, recall rate, and accuracy rate. The DC-LSTM algorithm has excellent performance in recall rate and accuracy rate. The recall rate of DC-LSTM is 0.95, and the accuracy rate is 0.93. The AUC curve, precision, recall, and accuracy of Video-STM were 0.71, 0.83, 0.61, and 0.73, respectively. The AUC curve, precision, recall, and accuracy of DC-LSTM are 0.72, 0.69, 0.68, and 0.80, respectively. Compared with other algorithms, DC-LSTM is more stable. To sum up, DC-LSTM has more advantages than single CNN and LSTM algorithms, and DC-LSTM also has more advantages than Video-STM and DC-LSTM mixed model. Therefore, DC-LSTM can help identify and analyze students' motion capture in physical education classroom,
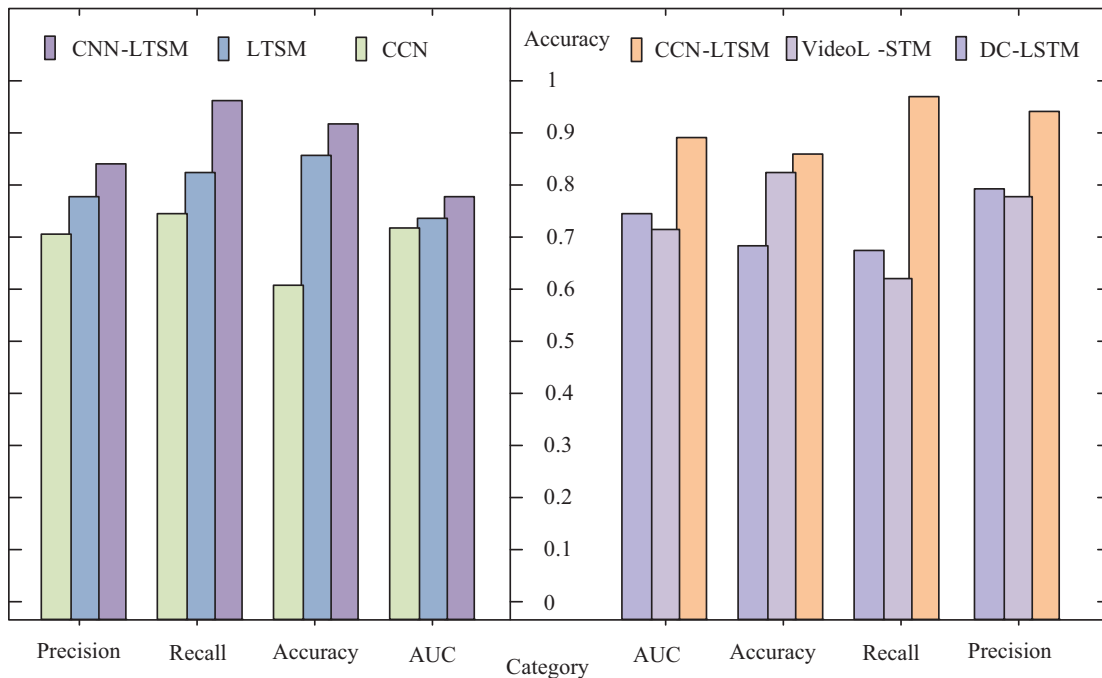

Fig. 10. Model comparison test.

in order to assist physical education and to improve students' physical quality.

V. CONCLUSION

In response to the low curriculum quality in middle school PEC, this study aimed to integrate CNN with long-term and short-term networks to obtain the CNN-LSTM algorithm. Based on this algorithm, a model was proposed to capture the physical activities of students in PEC. The performance comparison analysis of the proposed CNN-LSTM algorithm showed this algorithm's loss value was 0.045. It was lower than LSTM algorithm's 0.088, RCNN's 0.124, and CNN's 0.193. Its accuracy was 0.921, which was higher than LSTM algorithm's 0.863, RCNN's 0.815, and CNN's 0.796. Subsequently, in the practical verification of MCA, the proposed algorithm achieved recognition accuracy of 91.5%, 89.8%, 90.4%, 88.8%, and 90.1% for walking posture, jogging posture, running posture, boxing posture, and waving posture, respectively. In the analysis and comparison between DC-LSTM and the single algorithm, DC-LSTM had a higher recall rate than CNN and LSTM algorithms' 0.23 and 0.18, respectively, and a higher accuracy than CNN and LSTM algorithms' 0.31 and 0.04, respectively. In the comparison between DC-LSTM and hybrid model, the recall rate of DC-LSTM and DC-LSTM hybrid algorithm was about 0.34 and 0.25, respectively, and the accuracy was about 0.26 and 0.25, respectively, higher than these of the Video-STM and DC-LSTM hybrid algorithm. The above results showed that the MCA based on CNN-LSTM algorithm had a high accuracy and could promote the reform and development of middle school PEC. However, this research still had the problem of insufficient generality of the proposed algorithm. The subsequent research direction is to improve the generality of the algorithm, so as to improve the application field of the algorithm.

CONFLICT OF INTEREST STATEMENT

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

[1] T. A. Waziri and A. Ibrahim, "Discrete fix up limit model of a device unit," *J. Comput. Cognit. Eng.*, vol. 2, no. 2, pp. 163–167, 2022.

[2] D. Aikhuele, "Development of a statistical reliability-based model for the estimation and optimization of a spur gear system," *J. Comput. Cognit. Eng.*, vol. 2, no. 2, pp. 168–174, 2022.

[3] A. Sarkar, A. Biswas, and M. Kundu, "Development of q-Rung Orthopair Trapezoidal Fuzzy Einstein aggregation operators and their application in MCGDM problems," *J. Comput. Cognit. Eng.*, vol. 1, no. 3, pp. 109–121, 2022.

[4] B. Ma, Y. Fu, C. Wang, J. Li, and Y. Wang, "A high-performance insulators location scheme based on YOLOv4 deep learning network with GDIoU loss function," *IET Image Process.*, vol. 16, no. 4, pp. 1124–1134, 2022.

[5] K. Wan, X. Liu, J. Yu, X. Zhang, X. Du, and N. Guizani, "Compiler-based efficient CNN model construction for 5G edge devices," *IEEE Trans. Intell. Transport. Syst.*, vol. 22, no. 8, pp. 5261–5274, 2021.

[6] C. Zhang, F. Zhu, Y. Lv, P. Ye, and F. Wang, "MLRNN: taxi demand prediction based on multi-level deep learning and regional heterogeneity analysis," *IEEE Trans. Intell. Transport. Syst.*, vol. 23, no. 7, pp. 8412–8422, 2022.

[7] B. Batkhuu, J. Alikhanov, Y. Fang, S.H. Ko and G.S. Jo, "Transfer learning using multiple ConvNet layers activation features with principal component analysis for image classification," *J. Intell. Inf. Syst.*, vol. 24, no. 1, pp. 205–225, 2018.

[8] J. He, C. Li, P. Wang, Y. Liu, B. Liu, D. Xing, and Z. Wang, "Soliton molecule dynamics evolution prediction based on LSTM neural networks," *IEEE Photon. Technol. Lett.*, vol. 34, no. 3, pp. 193–196, 2022.

[9] X. Xu, H. Duan, Y. Guo, and Y. Deng, "A Cascade Adaboost and CNN algorithm for drogue detection in UAV autonomous aerial refueling," *Neurocomputing*, vol. 408, no. 3, pp. 121–134, 2020.

[10] Y. Fang, S.H. Ko and G.S. Jo, "Multi-scale region proposal network trained by multi-domain learning for visual object tracking," in Proc. of the 23rd International Conference on Neural Information Processing, Guangzhou, China, pp. 330–339, 2017.

[11] R. M. Kanko, E. K. Laende, G. Strutzenberger, M. Brown, and K. Deluzio, "Assessment of spatiotemporal gait parameters using a deep learning algorithm-based markerless motion capture system," *J. Biomech.*, vol. 122, no. 7, pp. 0021–0090, 2021.

[12] P. Gao, D. Zhao, and X. Chen, "Multi-dimensional data modelling of video image action recognition and motion capture in deep learning framework," *IET Image Process.*, vol. 14, no. 7, 1257–1264, 2020.

[13] S. Laaroussi, A. Baataoui, A. Halli, and K. Satori, "Dynamic mosaicking: combining A* algorithm with fractional Brownian motion for an optimal Seamline detection," *IET Image Process.*, vol. 14, no. 13, pp. 3169–3180, 2020.

[14] A. Chan, B. Coutts, E. Parent, and E. Lou, "Development and evaluation of CT-to-3D ultrasound image registration algorithm in vertebral phantoms for spine surgery," *Ann. Biomed. Eng.*, vol. 49, no. 1, pp. 310–321, 2020.

[15] C. Onks, D. Hall, T. Ridder, A. Idriss, J. Andrie, and R. Narayanan, "The accuracy and predictability of micro Doppler radar signature projection algorithm measuring functional movement in NCAA athletes – ScienceDirect," *Gait & Posture*, vol. 85, no. 2, pp. 96–102, 2021.

[16] M. Gil-Martín, R. San-Segundo, F. Fernández-Martínez, and J. Ferreiros-López, "Improving physical activity recognition using a new deep learning architecture and post-processing techniques," *Eng. Appl. Artif. Intell.*, vol. 92, no. 2, pp. 91–101, 2020.

[17] W. Lin, J. Leider, C. Shang, E. Hennessy, and F. Edd, "The association between state physical education laws and student physical activity – Sciencedirect," *Am. J. Prevent. Med.*, vol. 58, no. 3, pp. 436–445, 2020.

[18] P. Chonggao, "Simulation of student classroom behavior recognition based on cluster analysis and random forest algorithm," *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 2421–2431, 2021.

[19] P. Randhawa, V. Shanthagiri, and A. Kumar, "Violent activity recognition by E-textile sensors based on machine learning methods," *J. Intell. Fuzzy Syst.*, vol. 39, no. 6, pp. 8115–8123, 2020.

[20] B. Ferreira, P. M. Ferreira, G. Pinheiro, N. Figueiredo, F. Carvalho, and P. Menezes, "Deep learning approaches for workout repetition counting and validation," *Pattern Recogn. Lett.*, vol. 151, no. 9, pp. 259–266, 2021.

[21] W. Wei, A. N. You-wei, H. Zhan, D. Feng, and B. Cheng-xu. "The implementation of infrared image edge detection algorithm based on CNN on FPGA," *Guangzi Xuebao/Acta Photon. Sin.*, vol. 41, no. 11, pp. 1354–1358, 2012.

[22] L. Li et al., "Development of image-based wheat spike counter through a Faster R-CNN algorithm and application for genetic

studies," *J. Crops: English Edn.*, vol. 10, no. 5, pp. 1303–1311, 2022.

[23] K. Zhang and H. Shen, "Solder joint defect detection in the connectors using improved faster-RCNN algorithm," *Appl. Sci.*, vol. 11, no. 2, p. 576, 2021.

[24] Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. Lv, "Automatically designing CNN architectures using the genetic algorithm for image classification," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 1–15, 2020.

[25] D. Kang, D. Kang, and S. Ha, "Multi-bank on-chip memory management techniques for CNN accelerators," *IEEE Trans. Comput.*, vol. 71, no. 5, pp. 1181–1193, 2022.

[26] X. M. Long, Y. J. Chen, and J. Zhou, "Development of AR experiment on electric-thermal effect by open framework with simulation-based asset and user-defined input[C]," *Artif. Intell. Appl.*, vol. 1, no. 1, pp. 52–57, 2023.

[27] E. Nsugbe, "Toward a self-supervised architecture for semen quality prediction using environmental and lifestyle factors," *Artif. Intell. Appl.*, vol. 1, 1, pp. 35–42, 2023.

[28] S. Choudhuri, S. Adeniye, and A. Sen, "Distribution alignment using complement entropy objective and adaptive consensus-based label refinement for partial domain adaptation[C]," *Artif. Intell. Appl.*, vol. 1, no. 1, pp. 43–51, 2023.

[29] Z. Xu, F. Wu, X. Yang, and Y. Li, "Measurement of gas-oil two-phase flow patterns by using CNN algorithm based on dual ECT sensors with venturi tube. *Sensors*, vol. 20, no. 4, pp. 1200–1218, 2020.

[30] P. Wei, T. Jiang, H. Peng, H. Jin, and H. Huang, "Coffee flower identification using binarization algorithm based on convolutional neural network for digital images," *Plant Phenom.*, vol. 2020, no. 5, pp. 101–115, 2020.

[31] M. Z. Bellary, F Fameeza, and D. Musthafa, The MRI Knee pain classification using CNN algorithm and segmentation using clustering algorithm," *Turk. J. Comput. Math. Educ. (TURCOMAT)*, vol. 12, no. 10, pp. 306–315, 2021.

[32] F. Joy and V. Vijayakumar, "Multiple object detection in surveillance video with domain adaptive incremental fast RCNN algorithm," *Indian J. Comput. Sci. Eng.*, vol. 12, no. 4, pp. 1018–1026, 2021.