Application Effect of Human-Computer Interactive Gymnastic Sports Action Recognition System Based on PTP-CNN Algorithm

Yonge Ren[1], Keshuang Sun[2]*
Department of Physical Education, Jinzhong University, Jinzhong, 030600, China[1]
Department of Physical Education and Research, Fuzhou University, Fuzhou, 350108, China[2]

***Abstract*—With the rapid development of artificial intelligence technology, the recognition accuracy performance of traditional gymnastic sports action recognition system can no longer meet the needs of today's society. To address these problems, an improved action recognition algorithm combining Precision Time Protocal (PTP) and Convolutional Neural Networks (CNN) is proposed, and a human-computer interaction gymnastic action recognition system based on PTP-CNN algorithm is constructed. The performance test of the proposed PTP-CNN algorithm was conducted, and it was found that the accuracy of PTP-CNN algorithm was 92.8% and the recall rate was 95.2%, which was better than the comparison algorithm. The performance comparison experiments of the gymnastic action recognition system based on the PTP-CNN algorithm found that the recognition accuracy of the PTP-CNN gymnastic action recognition system was 96.3% and the running time was 3.4s, which was better than the other comparison systems. Comprehensive results can be found that the research proposed PTP-CNN recognition algorithm and improved gymnastic action recognition system can effectively improve the performance of traditional algorithms and models, which has practical application value and great application potential.**

***Keywords*—PTP; CNN; human-computer interaction; gymnastic sports; action recognition**

I. INTRODUCTION

In recent years, with the development of technology, the application of artificial intelligence technology in the field of sports has become more and more extensive [1]. Among them, the human-computer interactive gymnastics sports action recognition system is a new type of artificial intelligence technology that can achieve automatic recognition and analysis of gymnastic actions through machine learning and deep learning algorithms [2]. The application of this system can improve the efficiency and quality of gymnastics training, which is of great significance to promote the development of physical education and training. At present, many scholars at home and abroad have researched and practiced the human-computer interactive gymnastics sports movement recognition system [3]. Some of these studies are based on machine learning algorithms and deep learning algorithms for design and implementation [4]. At present, the most commonly used gymnastics action recognition is mainly based on the convolutional neural network (Convolutional Neural Networks, CNN). The CNN can learn the feature representation automatically, and it has a strong non-linear modeling capability. However, the traditional CNN mainly relies on the input of video frames and cannot make full use of the timing information of the action sequence. Moreover, the traditional CNN is prone to the problem of gradient disappearance or explosion when dealing with long time series. These problems limit the effectiveness of traditional sports systems in accuracy, real-time and personalized learning [5]. Accurate time protocol (PTP), as an accurate time synchronization protocol, has the advantages of high precision, high reliability and strong flexibility, which has a wide range in computer systems and networks. It can introduce time information in the action recognition task, and improve the accuracy and timing modeling ability of gymnastics movements of CNN by analyzing the forward and backward correlation in the action sequence [6]. At present, few studies have combined PTP with CNN and applied it in sports action recognition. Therefore, we propose to fuse PTP and CNN to build a PTP-CNN recognition algorithm and build a human-computer interactive gymnastic sports action recognition system based on it. It is expected to improve the efficiency and quality of gymnastics training and promote the development of physical education and training. Section I is the introduction to the article. The study describes the practical application of CNN and PTP in Section II. In Section III, PTP-CNN motion recognition algorithm and human-computer interactive gymnastics motion recognition system based on PTP-CNN are constructed. In Section IV, the action recognition algorithm and human-computer interaction action recognition system are tested. Results and discussion is given in Section V and Section VI concludes the paper.

II. REVIEW OF THE LITERATURE

With the continuous in-depth research on CNN algorithms by domestic and foreign scholars, various CNN improvement models have been proposed and applied in several fields. In order to improve the recognition accuracy of coffee flowers, Wei et al. combined CNN model with binarization algorithm, selected a certain number of positive and negative samples from the original digital images for network model training, initially extracted coffee flowers based on the trained CNN model, and then further optimized its boundary information using binarization algorithm, and experimentally verified that the accuracy of this method for coffee flower classification was 93.7%, which has practical Application significance [7]. Chowdary et al. proposed a measurement system based on improved convolutional neural network in order to improve the accuracy of mango leaf disease identification, and the proposed measurement was compared and analyzed with the system based on fuzzy algorithm, and the results showed that the accuracy of the proposed improved CNN-based system was 95.32, and this result indicated that the system could greatly improve the identification of mango leaf disease accuracy [8]. Joy and Vijayakumar found that the region-based convolutional neural network RCNN achieves good target detection accuracy but consumes more time on training and detection, so the proposed FAST RCNN algorithm with domain adaptive increments uses selective search to obtain the bounding box and feature extraction, thus overcoming the limitations of RCNN and improving the training and detection speed and accuracy [9]. Zhang's team proposed a deep learning model based on the combination of deep convolutional neural network and long and short-term memory network for the classification of arrhythmia intervals to address the problem that arrhythmias are difficult to diagnose accurately. Ten-fold cross-validation of the method showed that the average accuracy of the method was 99.06%, which is of great significance in clinical applications [10]. Zhao et al. For the problem of distortion and artifacts in lossy compressed video, a learning model with variable filter size residuals is proposed based on CNN algorithm, and the effectiveness of this model is measured using a combination based on color sensitivity, and extensive experimental results show that it has a better performance than existing methods in terms of efficiency improvement after video coding. [11].

With the rapid development of information technology, there are more and more methods applied in the field of action recognition. Rubin's team proposed a faster region based convolutional neural network structure to address the problem of difficult to accurately recognize real-time gestures, and tested its performance on a standard data set. Carvalho et al. proposed a multi-standard action-based human-robot interaction framework in order to improve the accuracy of socially assisted robot action recognition. [12]. Carvalho et al. tested the method offline and online, and the results showed that the accuracy of the method exceeded 96.7%, which is practical and can be used in educational proposals. [13]. Hu's team addressed the problem that current action recognition methods tend to ignore the reversibility of skeleton data in the temporal dimension [14]. Gao et al. proposed a new forward-inverse adaptive graph convolutional network for skeleton-based action recognition to address the problem that the current graph convolutional network models focus more on spatial information and ignore temporal information, and empirically analyzed the method. Gao et al. propose a unified attention model that integrates channel, space, and time, and the model is tested for performance and found to have the best performance compared to other similar action recognition methods [15].

The above studies fully illustrate that the CNN improvement model has been widely used in several fields, and there are also various methods applied in the field of action recognition. However, there are fewer studies combining PTP methods with CNN algorithms, so the study combines PTP methods with CNN algorithms to obtain PTP-CNN algorithms, and applies the improved algorithms to human-computer interactive gymnastic sports action recognition, expecting to improve the accuracy of gymnastic sports action recognition in this way and promote the further development of gymnastics course intelligence.

III. Construction of Human-computer Interaction Gymnastic Sports Action Recognition System based on PTP-CNN Algorithm

A. *CNN Action Recognition Algorithm Based on the Principle of PTP Protocol*

PTP protocol, also known as IEEE1588 protocol, is currently a mainstream time synchronization system, which is perfectly suitable for modern communication technology at the same time, but also well combined with computer hardware time equipment [16]. PTP can be placed in the computer network multiple clocks, and the master clock source and other clock sources in the form of telegrams time, to achieve accurate synchronization of network time. The PTP protocol clock type is divided into ordinary clock, boundary clock, end-to-end transparent clock and point-to-point transparent clock [17]. PTP synchronization principle implementation process is shown in Fig. 1.

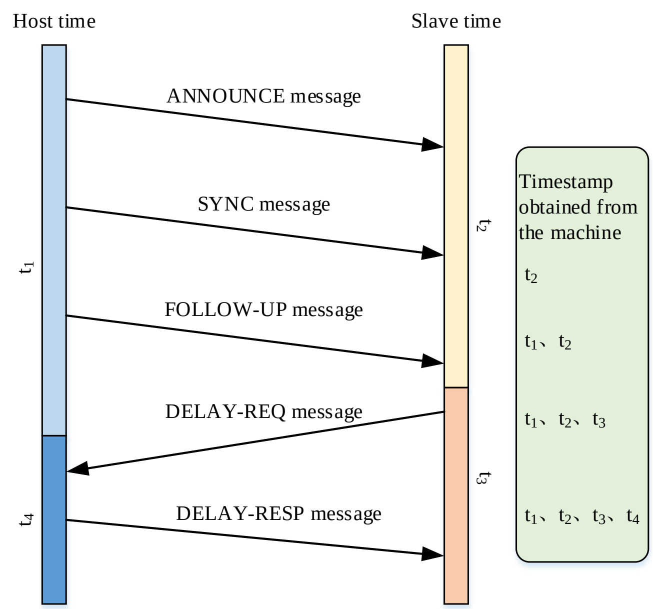


Fig. 1. Implementation process of PTP synchronization principle.

As it can be seen from Fig. 1, the PTP message flow is divided into four modules. First, the host sends an ANNOUNCE message to all devices in the slave, and the devices in the slave listening state will receive the ANNOUNCE message and set the host's clock source to the best master clock and set it to the uncalibrated state. Subsequently, the host sends a SYNC message to the slave and records the timestamp t_1, the slave receives it and records the timestamp t_2. The third module is the host sends a FOLLOW-UP message to the slave with the timestamp t_1 and the slave receives it and sends a DELAY-REQ message back to the host with the timestamp t_3. The host receives the

DELAY-REQ message and records the timestamp t_4. Finally, the host sends the DELAY-RESP message and the timestamp t_4 to the slave, and the slave finally gets a total of four timestamps. The slave obtains the link delay between the host and the slave by calculating the time deviation of the obtained timestamps from that of the host. The four timestamps t_1, t_2, t_1 and t_4 obtained by the slave are calculated as shown in Eq. (1).

$$\begin{cases} t_2 = t_1 + delay_{MS} + offset \\ t_4 = t_3 + delay_{SM} - offset \end{cases} \tag{1}$$

In Eq. (1), $offset$ is the time deviation from slave to host; $delay_{MS}$ is the link delay from host to slave; $delay_{SM}$ is the link delay from slave to host. When the values of $delay_{MS}$ and $delay_{SM}$ are the same, the link delay and time deviation are calculated as shown in Eq. (2).

$$\begin{cases} delay_{SM} = \dfrac{(t_2 - t_1) + (t_4 - t_3)}{2} \\ offset = t_2 - delay_{SM} \end{cases} \tag{2}$$

In Eq. (2), t_1, t_2, t_1 and t_4 are the time stamps obtained by the slave; $offset$ is the time deviation between the slave and the host; $delay_{MS}$ is the link delay from the host to the slave; $delay_{SM}$ is the link delay from the slave to the host. Due to the problem of link delay jitter in time synchronized networks in real networks. The study uses the second-order Kalman filtering algorithm to process the time deviation and link delay. The formula of first-order exponential smoothing filtering at this point is shown in Eq. (3).

$$y(n) = \alpha \cdot x(n) + (1-\alpha) \cdot y(n-1) \tag{3}$$

In Eq. (3), $y(n)$ denotes the value after the first n filtering and the initial value is 0; α is the coefficient of 0-1; $x(n)$ denotes the observed value of the first n. The cutoff frequency of the exponential smoothing filter is calculated as shown in Eq. (4).

$$\begin{cases} Y(Z) = \alpha \cdot X(Z) + (1-\alpha) \cdot Y(Z) \cdot z^{-1} \\ H(Z) = \dfrac{Y(Z)}{X(Z)} = \dfrac{\alpha}{1-(1-\alpha)\cdot z^{-1}} \end{cases} \tag{4}$$

In Eq. (4), $H(Z)$ denotes the system function; $Y(Z)$ denotes the discrete Z transformation of $y(n)$; α is the coefficient of 0-1; $Y(Z)$ denotes the discrete Z transformation of $x(n)$. The system frequency response is calculated as shown in Eq. (5).

$$H(e^{j\omega}) = H(Z)\big|_{z=e^{j\omega}} = \frac{\alpha}{1-1(1-\alpha)\cdot e^{-j\omega}} \tag{5}$$

In Eq. (5), $H(Z)$ denotes the system function; α is the coefficient of 0-1. The system amplitude and frequency response is shown in Eq. (6).

$$\left|H(e^{j\omega})\right| = \left|\frac{\alpha\left(\cos\omega + j\cdot\sin\omega\right)}{\cos\omega - (1-\alpha) + j\cdot\sin\omega}\right| = \frac{\alpha}{\sqrt{1+(1-\alpha)^2 + j\cdot\sin\omega}} \tag{6}$$

In Eq. (6), α is a factor of 0-1. $-3dB$ When the system cutoff frequency is calculated, the formula is shown in Eq. (7)

$$f = \frac{\arccos\left(1 - \dfrac{\alpha^2}{2(1-\alpha)}\right)}{2\pi} \tag{7}$$

In Eq. (7), α is a factor of 0-1. Since exponential smoothing filtering can effectively reduce the time deviation and thus improve the degree of system stability. The equation of control filtering at this point is shown in Eq. (8)

$$\begin{cases} l(n) = k\left(n\right)\cdot K_i + l(n-1) \\ m(n) = k(n)\cdot K_p + l(n) \end{cases} \tag{8}$$

In Eq. (8), $l(n)$ is the output value of the n integral control; $k(n)$ is the observed value of n; k_i is the coefficient of the integral control; $k(n)$ is the output value of the n exponential filter control; and k_p is the coefficient of the proportional control. The filtering process is shown in Eq. (9).

$$\begin{cases} filter_{es} = \alpha \cdot offset(n) + (1-\alpha)\cdot filter_{es}(n-1) \\ filter_l(n) = filter_{es}(n)\cdot K_l + filter_l(n-1) \\ filter_p(n) = filter_{es}(n)\cdot K_p + filter_p(n-1) \end{cases} \tag{9}$$

In Eq. (9), $filter_{es}$ is the exponential smoothing filter output value of n; α is the smoothing filter coefficient between 0 and 1; $offset(n)$ is the time deviation before the filter of n; $filter_i(n)$ is the integral control output value of n of the filter function; K_l is the integral control coefficient; $filter_p(n)$ is the proportional control output value of n of the filter function; K_p is the proportional control coefficient. After implementing the Kalman filter-based clock taming, the CNN action recognition algorithm based on the PTP principle needs to be constructed where the basic structure of CNN as shown in Fig. 2.

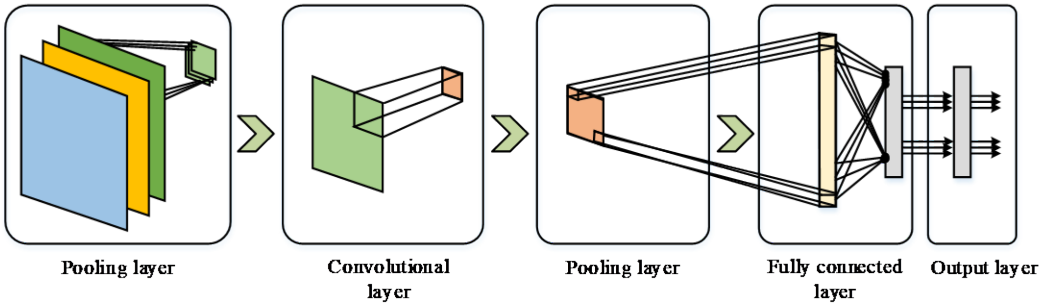


Fig. 2. CNN basic structure.

As shown in Fig. 2, CNN has five layers structure, which are input layer, convolutional layer, pooling layer, fully connected layer and output layer. Among them, the CNN convolutional layer is used for recognition and feature extraction of input data, including multiplication and addition of matrices, and its operation is closely related to the convolutional kernel [18-19]. The convolutional kernel is a feature extractor with sparse connectivity and weight sharing in the convolutional layer, and the convolutional kernel is trained to output a feature set that meets the needs [20]. The operation formula of the convolution layer is shown in Eq. (10).

$$x^l_j = f\left(\Sigma^n_{i\in M_j} x^{l-1}_i * k^l_{i,j} + b^i_j\right) \tag{10}$$

In Eq. (10), x^l and f are the output and activation function of the l layer, respectively; x^{l-1}_i is the output of the $l-1$ layer; k^l and b^l_j are the convolution kernel and offset term of the l layer; and M_j is the selected input feature set. The pooling layer divides the obtained feature set and reduces the dimensionality of the features to reduce the computational effort and enhance the robustness. The common operations of the pooling layer are mainly maximum pooling and mean pooling, and the two pooling methods are shown in Eq. (11).

$$a = \begin{bmatrix} 2 & 3 & 0 & 3 \\ 1 & 4 & 4 & 3 \\ 5 & 6 & 4 & 3 \\ 1 & 0 & 0 & 1 \end{bmatrix}, a_{\max} = \begin{bmatrix} 4 & 4 \\ 6 & 4 \end{bmatrix}, a_{ave} = \begin{bmatrix} 2.5 & 2.25 \\ 3 & 2 \end{bmatrix} \tag{11}$$

Finally, in order to prevent overfitting or underfitting of the model, reasonable optimization of the parameters is required. The study can use cross-entropy loss function and back propagation algorithm to calculate the error in fault multi-classification diagnosis. The formula of cross-entropy loss function is shown in Eq. (12).

$$H(p,q) = -\sum_x p(x)\log q(x) \tag{12}$$

In Eq. (12), $p(x)$ is the target distribution; $q(x)$ is the predictive distribution. The backpropagation algorithm is an algorithm for training a feedforward neural network for a given input pattern with a known classification using the chain derivation method. The backpropagation algorithm is the most common and effective method for training artificial neural network algorithms, and the essence is the error between the output and the target, as shown in Eq. (13).

$$\delta^{l-1} = \frac{\partial J}{\partial z^{l-1}} = \frac{\partial J}{\partial z^l}\frac{\partial z^l}{\partial z^{l-1}} = \delta^l \frac{\partial z^l}{\partial z^{l-1}}\frac{\partial a^{l-1}}{\partial z^{l-1}} \tag{13}$$

In Eq. (13), δ^l is the error of the objective function J to z^l. Finally, the PTP principle is fused with the CNN algorithm to construct the PTP-CNN recognition algorithm. the structure of the PTP-CNN recognition algorithm is shown in Fig. 3.

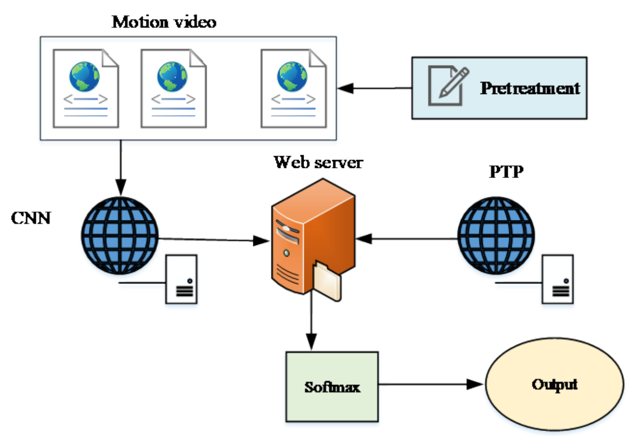


Fig. 3. PTP-CNN recognition algorithm structure.

Fig. 3 shows the structure of the PTP-CNN recognition algorithm, and the solid arrows in the figure represent the back-and-forth relationship between the modules in the PTP-CNN algorithm. As shown in Fig. 3, the PTP-CNN algorithm can extract the spatial features of video actions through CNN networks and fuse the convolutional features of different levels of CNN networks to enhance the feature representation. The temporal information in the video frames is modeled by PTP, and finally the classification results are obtained in using the class ware to recognize the human actions in gymnastics videos. To improve the recognition efficiency, the study first preprocesses the gymnastic video data with video frames, and adjusts the video frame size to (224*224) as the input. After completing the video pre-processing, the spatial features in the action video are extracted by CNN network and segmented into (224*224*3) size as input, the "3"in (224*224*3) indicates the channel dimension size. After that, the PTP-CNN algorithm fuses the shallow features with the deep features in the video through pooling and splicing operations to compensate for the missing feature information such as location contours. After several convolutions and pooling, the feature map with the output size of (7*7*2048) is pooled globally on average and expanded to (1*2048), using dropout to avoid overfitting of the results. Finally, the extracted feature vectors are input to the PTP time synchronization server to model the temporal information of the actions and the classification of the actions is achieved by a Softmax classifier.

B. Construction of Gymnastic Movement Recognition System Based on PTP-CNN Algorithm

After completing the construction of PTP-CNN recognition algorithm, the research will develop a gymnastic action recognition system. Since visual information is easily disturbed by external factors in real scenes, the accuracy and robustness of action recognition relying only on a single operational logic is very poor, so the research proposes a gymnastic action recognition system based on PTP-CNN algorithm. At the same time, in order to avoid the problem that the structure of the gymnastic action recognition system is confusing and difficult to expand, the research firstly designs the structure of the gymnastic action recognition system. The basic structure of the gymnastic action recognition system based on PTP-CNN algorithm is shown in Fig. 4.

As shown in Fig. 4, the gymnastic action recognition system proposed in the study takes reliability, practicality and scalability as design principles, and divides the system structure into four modules: web application layer, business logic layer, software development layer and basic function layer. The web application layer is the display page of the system to users, which includes user registration and login, video upload, action classification and recognition query functions. The business logic layer is the functional basis of the application layer, the main task is to manage the user database, complete the video pre-processing operation, and the video recognition and classification result analysis by PTP-CNN algorithm. The software development layer is the tool layer for system construction, including Pytorch framework, Python language, HTML/CSS front-end page development language and lightweight Flask framework, etc. The basic platform layer is the platform for the operation of the PTP-CNN gymnastic movement recognition system, and all the functions in the system cannot be realized without the support of hardware equipment and operation system. The hardware environment configuration of the gymnastic movement recognition system is AMD R5-4600H, 3.0GHz six-core twelve-thread processor, NVIDIA GeForce GTX1650Ti graphics card, 16GB running memory, 512GB solid state drive. The software environment is configured with Python language and Pytorch deep learning framework. the principle of PTP-CNN gymnastic action recognition system is shown in Fig. 5.

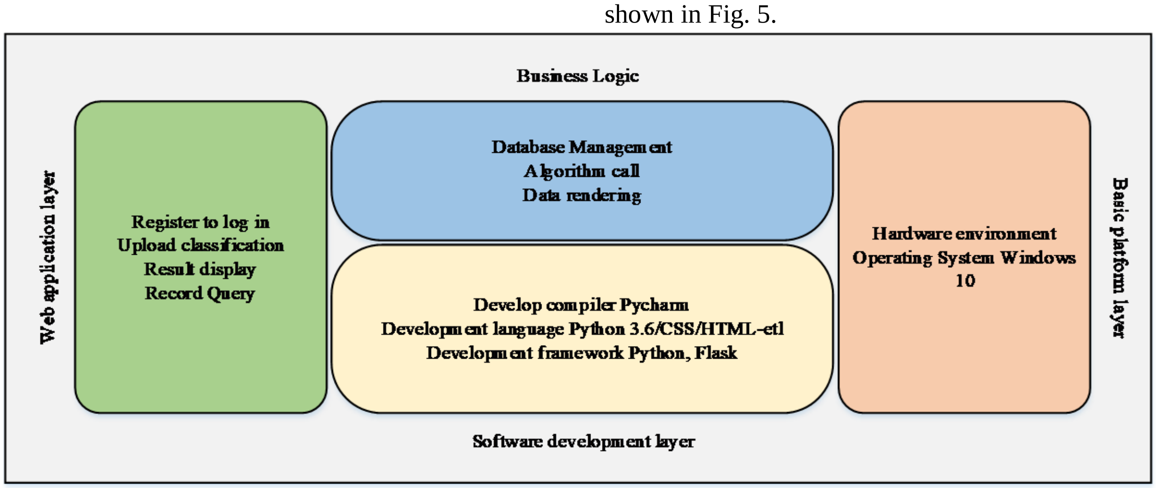


Fig. 4. Basic structure of gymnastic movement recognition system based on PTP-CNN algorithm.

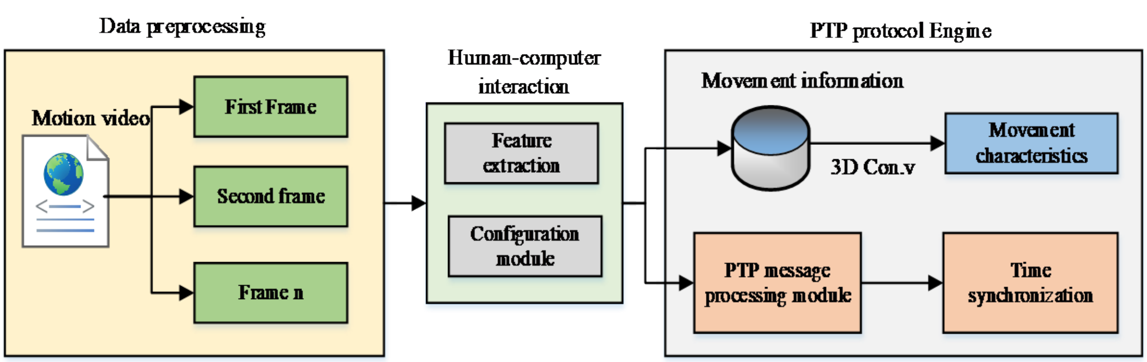


Fig. 5. Principle of PTP-CNN gymnastic movement recognition system.

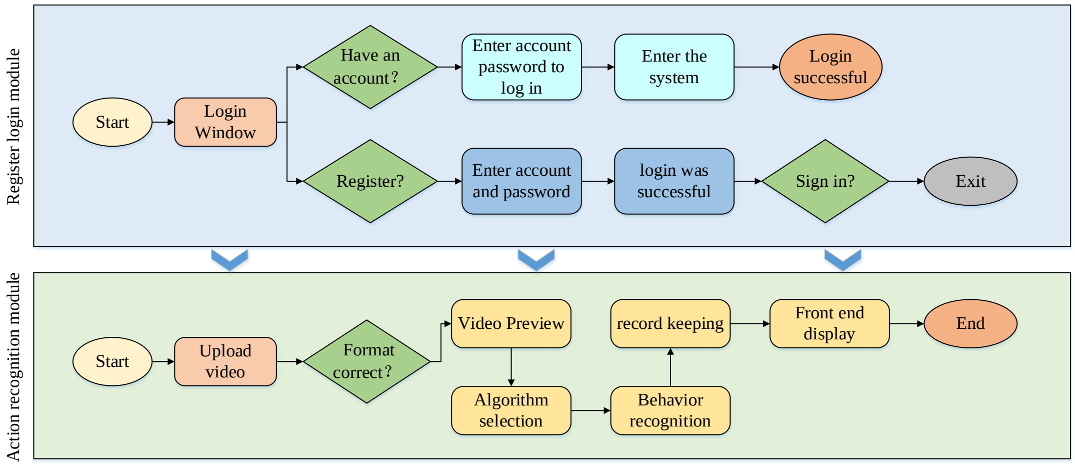


Fig. 6. The overall framework of PTP-CNN gymnastics motion recognition system.

As shown in Fig. 5, the proposed PTP-CNN gymnastic action recognition system consists of three modules, namely, the data pre-processing module, the human-computer interaction module and the PTP protocol engine module. In the data pre-processing module, the system decomposes the video in the dataset into continuous frames and feeds them into the HCI module. The principle of HCI module is to finish the feature recognition and feature extraction of gymnastics video by CNN algorithm. Finally, the PTP protocol engine module takes the video action information and different action features extracted by the CNN algorithm, models and fuses them in 3D, and outputs the gymnastic action feature recognition results. The overall framework of PTP-CNN gymnastics action recognition system is shown in Fig. 6.

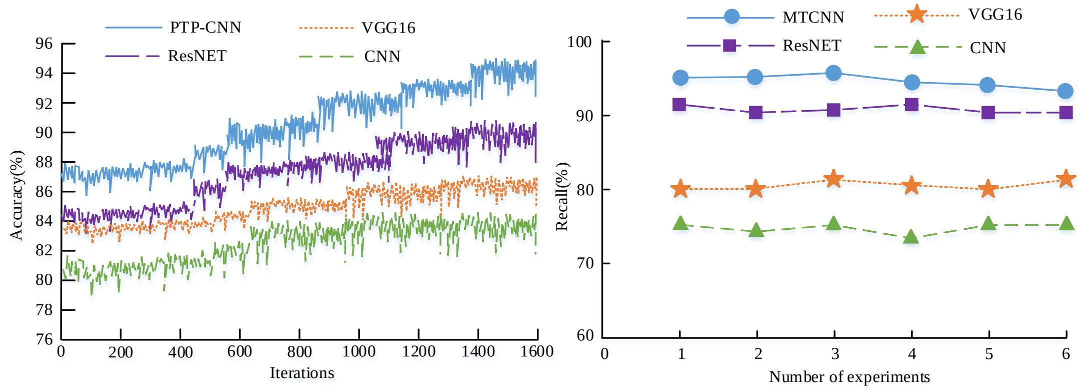
As shown in Fig. 6, the PTP-CNN gymnastic action recognition system is mainly divided into the user's login and registration module, and the action recognition module of the underlying logic of the system. In the login and registration module, after the user enters the system login window, the system takes the user's username at the time of registration as the login account, and judges the submitted username when the user submits the registration information. If the user name is recognized to exist in the database, the user registration fails, otherwise the registration is successful, at this time the PTP-CNN gymnastics action recognition system generates a unique user number primary key for the subsequent query operation of the user. After the registration is completed, users can enter the system by entering the correct account password, at which time they can upload the gymnastics video. Since the PTP-CNN gymnastics action recognition system only supports video uploads in avi and mp4 formats, the system will identify the format of the video uploaded by the user. If the format is correct, it will enter the video preview stage; if the video format is wrong, it will return to the video upload stage. After entering the video recognition stage, the system will call the PTP-CNN algorithm in the background for action recognition and write the recognition result to the database history. After finishing the gymnastic action classification recognition, the background will package the recognition results into JSON format data to return to the front-end, and the results will be displayed in the user interface.

IV. Empirical Analysis of Human-Computer Interactive Gymnastic Sports Action Recognition System Based on PTP-CNN Algorithm

A. *Analysis of the Effectiveness of PTP-CNN Recognition Algorithm*

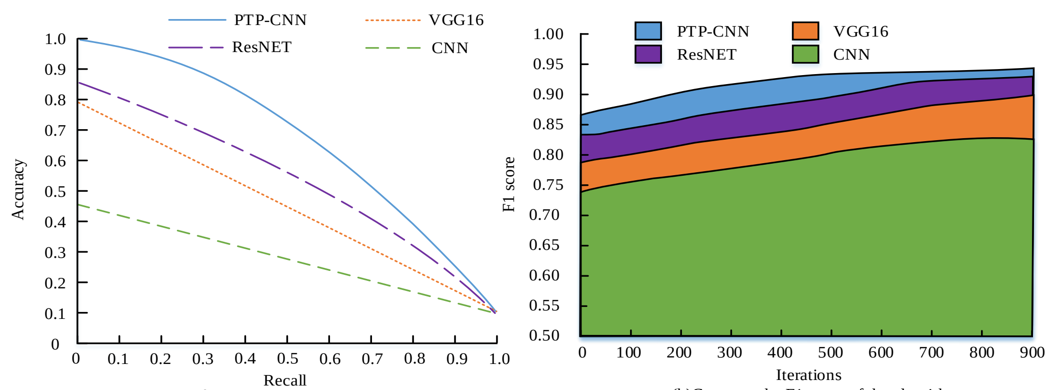
To verify the effectiveness of the PTP-CNN recognition algorithm for gymnastic sports action recognition, the study uses the public dataset UTKinect-Action3D to validate the effectiveness of the PTP-CNN recognition algorithm. the UTKinect-Action3D dataset contains activities corresponding to the gymnastic action recognition system, such as finishing exercises, stretching exercises, chest expansion exercises, full body movement, body rotation movement and jumping movement. The PTP-CNN recognition algorithms are compared and analyzed by the recognition accuracy, recall, F1 value, verification loss value (val-loss), and verification accurate values (val-acc) of these six corresponding actions. (Visual Geometry Group Network-16, VGG16), Residual Network (ResNET) and Dynamic Time Warping (DTW) algorithms. The recognition accuracy and recall curves of the compared algorithms are shown in Fig. 7.

Fig. 7(a) shows the recognition accuracy curves of the compared algorithms. From Fig. 7(a), it can be seen that the recognition accuracy of each algorithm increases with the number of iterations, and the PTP-CNN algorithm proposed in the study has an overall higher accuracy than the other algorithms, with an accuracy rate of up to 94.3% and an average accuracy rate of 92.8%. Fig. 7(b) shows the recall curves of the compared algorithms. From Fig. 7(b), it can be seen that the recall rate of each algorithm is smooth and does not change with the number of experiments, among which the PTP-CNN algorithm proposed in the study has a higher recall rate than the other algorithms, and its average recall rate is 95.2%. From the above results, it is clear that the accuracy performance and recall performance of the PTP-CNN algorithm proposed in the study are better than the other algorithms. Fig. (8) shows the accuracy-recall rate curves and F1 value comparison results of each compared algorithm.



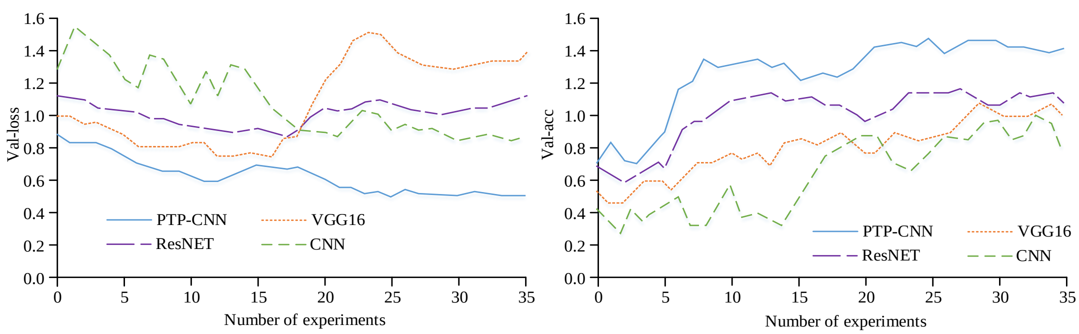
(a)Accuracy curve of comparison algorithm (b)Comparison algorithm recall curve

Fig. 7. Accuracy and recall curves of each algorithm.



(a)The first comparative experiment (b)Compare the F1 score of the algorithm

Fig. 8. Accuracy-recall and F1 score curves of each algorithms.



(a)Val-loss curves of various comparison algorithms (b)Val-acc curves of various comparison algorithms

Fig. 9. Validation losses and validation accuracy of each algorithm.

Fig. 8(a) shows the accuracy-recall curves of the compared algorithms. From Fig. 8(a), it can be seen that the accuracy-recall curve of the proposed PTP-CNN algorithm has the largest area under the line of 0.81 compared with other algorithms, which is 0.5 larger than the accuracy-recall curve of CNN. Fig. 8(b) shows the F1 values of each comparison algorithm. From Fig. 8(b), it can be seen that the proposed PTP-CNN algorithm has the largest F1 value of 0.93 compared to the other compared algorithms, which is 0.08 higher than the F1 value of CNN. In summary, it can be seen that the proposed PTP-CNN algorithm has the best performance in terms of accuracy-recall rate and F1 value performance. Fig. 9 shows the val-loss and val-acc values of the compared algorithms.

Fig. 9(a) shows the val-loss curves of each comparison algorithm. From Fig. 9(a), it can be seen that the val-loss curve of the proposed PTP-CNN algorithm has the lowest overall val-loss curve and the smallest fluctuation compared to the other comparison algorithms, with an average val-loss value of 0.72 and a fluctuation of 0.31. Fig. 9(b) shows the val-acc values of each comparison algorithm. From Fig. 8(b), it can be seen that the PTP-CNN algorithm proposed in the study has the largest val-acc value of up to 1.51 compared with the other comparison algorithms, which is 0.58 higher than the highest val-acc value of CNN. In summary, the results show that the PTP-CNN algorithm proposed in the study has the best performance in terms of the performance of verification loss and verification accuracy.

A. Comparison Experiment of Gymnastic Movement Recognition System Based on PTP-CNN Algorithm

To test the recognition performance of the PTP-CNN algorithm-based gymnastic movement recognition system, the study conducts a comparative performance analysis of the system. The study tests the performance of the system by comparing the recognition accuracy, recognition error and system running time, etc. The comparison system is a gymnastic action recognition system based on ResNET, VGG16 and CNN algorithm. The system test platform is composed of KinectV1.0, Windows10, VisualStudio and Unity. The recognition accuracy of each system is shown in Fig. 10.

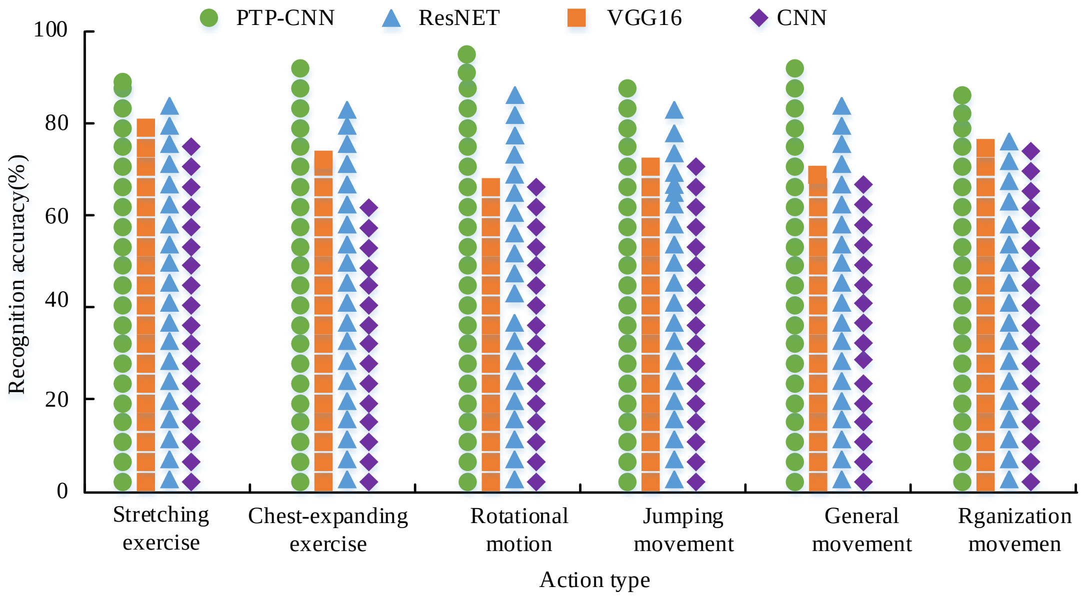


Fig. 10. Identification accuracy of each system.

Fig. 10 shows the recognition accuracy of the action recognition system based on PTP-CNN, ResNET, VGG16 and CNN algorithms for stretching, chest expansion, body turn, jumping, full-body and finishing movements. As shown in Fig. 10, the recognition accuracy of the PTP-CNN algorithm-based gymnastics action recognition system is higher than other comparison systems in six movements as a whole, among which the system has the lowest recognition accuracy of 93.4% for finishing movements and the highest recognition accuracy of 98.5% for full-body movements. The average recognition accuracy of the PTP-CNN algorithm-based gymnastic movement recognition system is 96.3%, which is 10.1% more accurate than that of the ResNET algorithm-based recognition system. Summing up the results, it can be concluded that the gymnastic movement recognition system based on PTP-CNN algorithm proposed in the study has the best performance in terms of movement recognition accuracy. The confusion matrix of the PTP-CNN algorithm-based gymnastic action recognition system is shown in Fig. 11 when the recognition results of the actions are compared with the actual actions.

Fig. 11 shows the confusion matrix of the gymnastic action recognition system based on PTP-CNN algorithm. From Fig. 11, it can be seen that the recognition accuracy of the gymnastic action recognition system proposed in the study is high on all six actions, indicating that the system has excellent recognition and classification ability on and distinguished actions, which has practical use value. To further verify the practical use performance of the PTP-CNN algorithm-based gymnastic action recognition system, the study conducted empirical experiments on the system to analyze its system recognition error and system computing speed, and the results of the empirical experiments are shown in Fig. 12.

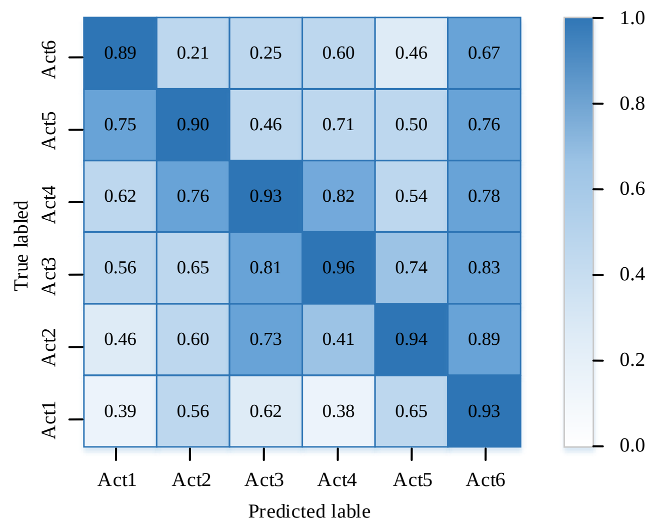


Fig. 11. PTP-CNN gymnastics recognition system confusion matrix.

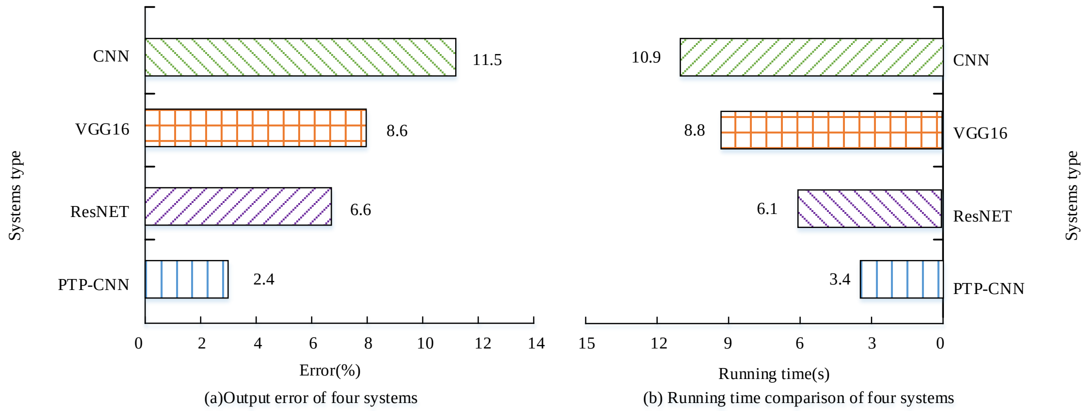


Fig. 12. Identification error and calculation time of each system.

Fig. 12(a) shows the recognition errors of gymnastic movements for each comparison system. From Fig. 12(a), it can be seen that among the four comparison systems, the system with the largest recognition error is the CNN algorithm-based gymnastic movement recognition system, whose recognition error is 11.5%. And the recognition system based on PTP-CNN algorithm has the smallest recognition error among the comparison systems, which is 2.4% and 9.1% lower than the gymnastic action recognition system based on CNN algorithm. Fig. 12(b) shows the operation speed of each comparison system. From Fig. 12(b), it can be seen that among the four comparison systems, the longest running speed is the CNN algorithm-based gymnastic movement recognition system, which runs for 10.9 s. The PTP-CNN algorithm-based recognition system has the shortest running time among the comparison systems, which is 3.4 s and 7.5 s shorter than the CNN algorithm-based gymnastic movement recognition system. Gymnastic movement recognition system has the best performance in terms of recognition error and running speed, and has significantly improved the performance compared with the traditional gymnastic movement recognition system.

V. Results and Discussion

It is found that the PTP-CNN motion recognition system can accurately identify the stretching, chest expansion, body turning, body jumping, whole body movement and finishing movement in gymnastics, with the accuracy rate of 96.3%, which is better than the traditional CNN motion recognition system. The findings are consistent with Yu et al. [21], which improved the accuracy of action classification and were applied to EMG control. In addition, the study also found that PTP-CNN human-computer interactive gymnastics motion recognition system has a short running time and has practical application value. Similar to the study of Majd et al. [22]., can be applied in the fields such as video surveillance. In addition, the action recognition system was applied to the path planning of the traffic system by Chen et al. [23]. Therefore, the proposed human-computer interaction gymnastics motion recognition system has good application potential in the fields of medicine, video surveillance, traffic and sports. In the medical field, the system can be used for rehabilitation training and evaluation, helping doctors to monitor patients' exercise recovery and provide targeted rehabilitation programs. In the field of video surveillance, the system can be used to monitor and identify human movements in real time, and help security personnel quickly detect abnormal or criminal behaviors. In the field of traffic, the system can be used to identify the driver's action and posture, monitor the driver's fatigue driving situation, remind the driver to pay attention to safety, so as to reduce the occurrence of traffic accidents. Most importantly, in the field of sports, the system can be applied in training and competition to help coaches and athletes analyze and improve movement skills, and improve training results and competition performance. In addition, the system can be used to evaluate the performance of the players and provide objective basis for the judges. In conclusion, the human-computer interactive gymnastics motion recognition system has wide application potential to play an important role in medicine, video surveillance, transportation and sports.

VI. Conclusion

In order to improve the recognition accuracy and operation speed of the traditional gymnastic action recognition system, the study proposes an action recognition algorithm that integrates the PTP principle and CNN algorithm, and builds a gymnastic action recognition system based on the PTP-CNN algorithm based on it. The performance tests of the proposed PTP-CNN algorithm and the improved gymnastic action recognition system are conducted. The results show that the PTP-CNN algorithm has the highest accuracy of 94.3%, the average accuracy of 92.8%, and the recall rate of 95.2%, which are better than the rest of the comparison algorithms in terms of accuracy performance and recall rate performance. In addition, the study also conducted performance comparison experiments on the improved gymnastic movement recognition system based on the PTP-CNN algorithm. The results show that the recognition accuracy of PTP-CNN gymnastic action recognition system is 96.3%, which is better than the other comparison systems. In addition, it is found that the average running time of PTP-CNN gymnastic action recognition system is 3.4s, which is lower than other comparison systems. The above results can be found that the proposed PTP-CNN recognition algorithm and gymnastic movement recognition system are better than the comparison algorithm and system in terms of recognition accuracy and running speed, and have practical application value. However, the study also has some limitations, compared to the gymnastics posture, including the extension, distortion and rotation of the body. This attitude change poses certain challenges to the accuracy and robustness of the algorithm. Future studies can explore more effective posture feature extraction methods according to the large posture changes in gymnastics movements, and improve the algorithm's ability to identify and model posture changes.

ACKNOWLEDGMENT

The research is supported by: 2022 Shanxi Provincial College Philosophy and Social Sciences Project Establishment, Research on the Current Situation and Construction Paths of " Ideological and Political Courses & quot; in Public Courses in Colleges and Universities in Shanxi Province, (No. 2022W135).

REFERENCES

[1] Qian J. Research on Artificial Intelligence Technology of Virtual Reality Teaching Method in Digital Media Art Creation. Journal of Internet Technology, 2022, 23(1):125-132.

[2] Modi N, Singh J. Role of Eye Tracking in Human Computer Interaction. ECS transactions, 2022, 107(1):8211-8218.

[3] Xu W. From Automation to Autonomy and Autonomous Vehicles: Challenges and Opportunities for Human-Computer Interaction. interactions, 2021, 28(1):48-53.

[4] Oslund S, Washington C, So A, Chen T, Ji H. Multiview Robust Adversarial Stickers for Arbitrary Objects in the Physical World. Journal of Computational and Cognitive Engineering, 2022, 1(4): 152-158.

[5] Bae J, Lee D H. PTP Tracking Scheme for Indoor Surveillance Vehicle by Dual BLACM With Hall Sensor. IEEE Transactions on Industry Applications, 2022, 58(4):5238-5247.

[6] Nimrah S, Saifullah S. Context-Free Word Importance Scores for Attacking Neural Networks. Journal of Computational and Cognitive Engineering, 2022, 1(4): 187-192.

[7] Wei P, Jiang T, Peng H, Jin H, Huang H. Coffee Flower Identim fication Using Binarization Algorithm Based on Convolutional Neural Network for Digital Images. Plant Phenomics, 2020, 5:101-115.

[8] Chowdary M S, Puviarasi R. Accuracy Improvement in Disease Identification of Mango Leaf using CNN Algorithm Compared with Fuzzy Algorithm. ECS transactions, 2022, 107(1):11889-11903.

[9] Joy F, Vijayakumar V. Multiple Object Detection In Surveillance Video With Domain Adaptive Incremental Fast Rcnn Algorithm. Indian Journal of Computer Science and Engineering, 2021, 12(4):1018-1026.

[10] Zhang P, Hang Y, Ye X, Guan P, Hu W. A United CNN-LSTM Algorithm Combining RR Wave Signals to Detect Arrhythmia in the 5G-Enabled Medical Internet of Things. IEEE Internet of Things Journal, 2021, 9(16):14563-14571.

[11] Zhao H, He M, Teng G, Shang X, Wang G, Feng Y. A CNN-Based Post-Processing Algorithm for Video Coding Efficiency Improvement. IEEE Access, 2020, 8(1):920-929.

[12] Rubin B S, Kumar V S. An Efficient Inception V2 based Deep Convolutional Neural Network for Real-Time Hand Action Recognition. IET Image Processing, 2019, 14(4):688-696.

[13] Carvalho K, VT Basílio, Brando A S. Action recognition for educational proposals applying concepts of Social Assistive Robotics. Cognitive Systems Research, 2022, 71:1-8.

[14] Hu Z, Pan Z, Wang Q, Yu, L, Fei S. Forward-reverse adaptive graph convolutional networks for skeleton-based action recognition. Neurocomputing, 2022, 492(1):624-636.

[15] Gao B K, Dong L, Bi H B, Bi Y Z. Focus on temporal graph convolutional networks with unified attention for skeleton-based action recognition. Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2022, 52(5):5608-5616.

[16] Ren Y, Zhang R, Feng Z, Ke C, Yao S, Tang C, Lin L, Ye Y. Macrocephatriolides A and B: Two Guaianolide Trimers fromAinsliaea macrocephalaas PTP1B Inhibitors and Insulin Sensitizers. The Journal of organic chemistry, 2021, 24(21):17782-17789.

[17] J García-Marín, Griera M, R Alajarín, M Rodríguez-Puyol, D Rodríguez-Puyol, Vaquero J. A computer-driven scaffold-hopping approach generating new PTP1B inhibitors from the pyrrolo[1,2-a] quinoxaline core. ChemMedChem, 2021, 18(16):2895-2906.

[18] Chen B, Xingwang J U, Gao Y, Wang J. A Quaternion Two-Stream R-CNN Network for Pixel-Level Color Image Splicing Localization. Chinese Journal of Electronics, 2021, 30(6):1069-1079.

[19] Joardar B K, Doppa J R, Pande P P, Li H, Chakrabarty K. AccuReD: High Accuracy Training of CNNs on ReRAM/GPU Heterogeneous 3D Architecture. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2020, 5(40):971-984.

[20] Sun K, Zhang J, Liu J, Yu R, Song Z. DRCNN: Dynamic Routing Convolutional Neural Network for Multi-View 3D Object Recognition. IEEE Transactions on Image Processing, 2020, 30(12):868-877.

[21] Yu B, Zhang X, Wu L. A Novel Postprocessing Method for Robust Myoelectric Pattern-Recognition Control Through Movement Pattern Transition Detection.IEEE transactions on human-machine systems, 2020, 50(1):32-41.

[22] Majd M, Safabakhsh R. A motion-aware ConvLSTM network for action recognition. Applied Intelligence, 2019, 49(7):1-7.

[23] Chen H Y, Huang P H, Fu L C. Social crowd navigation of a mobile robot based on human trajectory prediction and hybrid sensing. Autonomous robots, 2023, 47(4):339-351.