Jumping Action Recognition for Figure Skating Video in IoT Using Improved Deep Reinforcement Learning

Yu Liu

Physical Education Department, Harbin Engineering University, Harbin 150001, China;
e-mail: liuyu6666666@163.com

Ning Zhou

Winter Olympic College, Harbin Sport University, Harbin 150008, China; e-mail: zhouning@hrbipe.edu.cn

Corresponding author: zhouning@hrbipe.edu.cn

Figure skating video jumping action is a complex combination action, which is difficult to recognize, and the recognition of jumping action can correct athletes' technical errors, which is of great significance to improve athletes' performance. Due to the recognition effect of figure skating video jumping action recognition algorithm is poor, we propose a figure skating video jumping action recognition algorithm using improved deep reinforcement learning in Internet of things (IoT). First, IoT technology is used to collect the figure skating video, the figure skating video target is detected, the human bone point features through the feature extraction network is obtained, and centralized processing is performed to complete the optimization of the extraction results. Second, the shallow STGCN network is improved to the DSTG dense connection network structure, based on which an improved deep reinforcement learning action recognition model is constructed, and the action recognition results are output through the deep network structure. Finally, a confidence fusion scheme is established to determine the final jumping action recognition result through the confidence is established. The results show that this paper effectively improves the accuracy of figure skating video jumping action recognition results, and the recognition quality is higher. It can be widely used in the field of figure skating action recognition, to improve the training effect of athletes.

KEYWORDS: Action recognition, Figure skating, Improve deep reinforcement learning, Internet of things, Figure Skating Video, Dense connection network.

1. Introduction

As one of the core technologies of the Internet of things (IoT), human action recognition has attracted much attention in recent years and has begun to play a role in many fields [15]. In particular, in the field of figure skating, jumping is a core technical action, which directly determines the performance of athletes [4, 14]. Therefore, whether in teaching or training of figure skating, it is necessary to accurately recognize athletes' jumping actions through videos, so as to improve their action quality [17]. Most of the current action recognition technologies are aimed at simple actions such as standing and raising hands, which can be recognized according to the key markers in the video [23]. However, jumping is a complex compound action [12]. Thus, it is difficult to accurately obtain the jumping action recognition results from the figure skating video by using the traditional action recognition algorithm [8]. Mazurkiewicz [10] analyzed the data of single person figure skater's single person and maximum total rotation vertical jump under the condition of de-icing. The piezoelectric sensor platform and Bio-Ware software were used to calculate the ground reaction force. In combination with the recognition results of the ground reaction force, Vicon T series action capture technology was used to recognize the jumping action in figure skating. However, the error of this algorithm is huge. In Delisle-Houde et al. [1], based on the skating videos of several athletes, the athletes were required to perform standing long jump, vertical long jump, Axel jump, Lutz hook jump, toe-loop jump, triple flip jump, loop jump, etc. to obtain video data. Pearson correlation was used to analyze the relationship between agility and action completion, to obtain the features of athletes' jumping actions and recognize jumping actions. However, the average recognition rate of this algorithm is low. Wang and Tong [16] analyzed the body changes of advanced dance movements based on deep learning algorithm and IoT to promote the practical application of biological image visualization technology. Deep learning and IoT edge computing are used to build the network structure of the dance generation model. The experimental results show that the deep learning based Internet of things improves the efficiency of resource allocation and effectively reduces the server processing time. However, the recognition

accuracy of this algorithm is low. Zhai et al. [22] used wavelet coefficient transform to get the noise threshold of moving video image. Then the wavelet coefficient matrix was constructed based on decomposing the image multi-resolution, and the image noise was removed according to the threshold. On this basis, the Gaussian function in the canny method was introduced, and the edge coordinates of the moving target were obtained through convolution operation. According to the Murkowski distance, the image rotation and translation matrix were analyzed to obtain the action feature vector. A hidden Markov model was constructed to segment the background and the target according to the feature sequence and the target edge coordinates, to realize the jumping action recognition. However, this algorithm has the problem of poor recognition effect. Therefore, it is difficult to be applied to practice. In Nguyen et al. [13], based on the action video, the human action data in the three-dimensional coordinate system were preprocessed and normalized by calculating the deviation from the average coordinate. The data was transformed into vectors in conformal geometric algebraic space, and its dimension was reduced to return eigenvectors. The recurrent neural network model was used to train the feature vector of human jumping action, and the result of feature vector extraction was combined to recognize human jumping action. However, in practical application, it is found that this algorithm has the problem of poor recognition effect.

However, the above methods have low Average Precision (AP) value, Area Under Curve (AUC), true positive rate, intersection over union (IoU) and F-measure value in the process of identifying jumping action in figure skating videos. Therefore, this paper proposes an algorithm to recognize the jumping action in figure skating videos based on improved deep reinforcement learning in the IoT. The contributions of this paper are as follows: (1) The IoT technology is used to collect video samples of figure skating, which can ensure the effect and quality of data collection from figure skating videos and provide an important data basis for subsequent jumping action recognition. (2) The shallow STGCN network is improved to the DSTG dense connection network structure, so as to build an improved deep reinforcement learning action recognition model

to ensure the accuracy and efficiency of the recognition results. (3) Figure Skating data set, MSR Action 3D data set, UCF 101 data set and HMDB51 data set are selected for experimental tests to ensure the authenticity and reliability of experimental results and improve the reliability of experiments.

2. Methodology

2.1. Figure Skating Video Acquisition and Target Detection Based on IoT

In this paper, the IoT is mainly to collect figure skating videos, to detect the target in the video, and lay a solid foundation for the subsequent jumping action recognition from figure skating videos. Figure 1 shows the structure of figure skating video acquisition based on IoT.

According to Figure 1. The capture architecture of figure skating videos proposed algorithm draws on the general IoT architecture, including the business layer, network layer and perception layer. Among them, the perception layer is mainly composed of Radio Frequency Identification (RFID), sensors, Global Positioning System (GPS) and intelligent terminals, which are responsible for the data collection from figure skating videos. The data collected by the sensing layer is transmitted to the service layer through the mobile communication network, computer network and wireless sensor network in the network layer to complete the data exchange. The collected data from figure skating videos are analyzed in the business layer to complete the data acquisition from figure skating

Figure 1
Figure Skating video acquisition architecture based on IoT
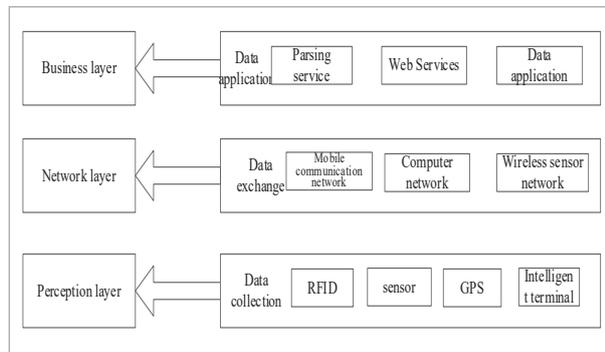
videos, which is used as the design basis of the action recognition algorithm. First, the key image frames are selected for the video collected by the IoT, and the video target detection method is designed. Then the space-time condition information of the image block is obtained, and the marking matrix is used to detect the moving target.

$$A(m,n) = \begin{cases} 2, & V(m,n) \geq r_V \\ 1, & C(m,n) \geq r_C \\ 0, & others \end{cases} \quad (1)$$

where (m, n) refers to the human body coordinate position in the figure skating videos, A is the marked matrix. V refers to image block difference information. C refers to space time condition information of image block, r_C and r_V are the classification thresholds of V and C, respectively.

After image target extraction, it is necessary to track and calculate the gait contour as the basis of recognition algorithm design. Through 11 gait contour points, the force features of jumping actions in figure skating are described [9]. In the pattern video, there is periodicity in the appearance of gait contour points. To track the athlete's action gait contour, the contour period calculation formula is proposed as follows:

$$t = \begin{cases} \sin(t + \lambda), & -\lambda + E_1 \leq t \leq E_1 + \dfrac{E}{2} - \lambda \\ 0, & others \\ \sin(t - 2E_1 - \dfrac{E}{2} + \lambda), & -\lambda + 2E_1 \leq t \leq 2E_1 + E - \lambda \end{cases} \quad (2)$$

where t refers to the gait contour period of jumping action. E stands for the total time of the gait contour period. λ refers to the phase angle, and E_1 is the support time of key gait contour points during jumping.

In addition, according to the video collected by the IoT, the action space is analyzed, the action step size, action speed and other data are obtained, and the displacement model is established based on the real-time displacement change data. Then a grid region is designed, and the output results of gait period model and displacement model are displayed in the tracking grid region as the basis for subsequent feature extraction and action recognition.

2.2. Feature Extraction and Optimization of Bone Point

Based on the above figure skating video acquisition and target detection results, human bone point features are obtained through the feature extraction network [20], and the feature extraction results are optimized. A feature extraction network consisting of two branches is designed to synchronously extract regional confidence map features and regional correlation field features [2]. Then all feature maps are fused to obtain feature extraction results

$$U^\varepsilon = \rho^\varepsilon(G, U^{\varepsilon-1}, K^{\varepsilon-1}), \forall \varepsilon \geq 2, \qquad (3)$$

where ε refers to the feature extraction time, U is regional confidence map, G refers to input features, and ρ is feature extraction network.

Figure skating video contains more audience bone points, which has a negative impact on the feature extraction results. Therefore, in this paper, data cleaning is carried out after the feature extraction of bone points. For each image frame, the bone points of two people are extracted to ensure that these two groups of bone points have the highest confidence [6], and the feature information of the remaining bone points is proposed. Since the protagonist of the figure skating video is an athlete, the group with the largest area is selected as the feature extraction result of the two groups of bone points. In addition, the centralized processing mode [19] is adopted to reduce the fluctuation of bone point data, as shown in Figure 2.

According to the data in Figure 2. Node 1 represents clavicle; Nodes 2, 3 and 4 represent the right scapula, ulna and wrist, respectively; Nodes 5, 6 and 7 represent the left scapula, ulna and wrist, respectively; Node 8 represents the coccyx; Nodes 12 and 9 represent the pubic bones on the left and right sides, respectively; Nodes 13 and 10 represent the left and right kneecaps, respectively; 10,24, 22 and 23, respectively represent the fibula, tarsal, metatarsal and metatarsal bones on the right side; Nodes 14, 21, 19 and 20 represent the left fibula, tarsal, metatarsal and metatarsal bones.

The specific process of implementing the centralized treatment scheme is as follows:

$$y = d \times x \qquad (4)$$

$$d = \delta \times \tau \qquad (5)$$

Figure 2
Schematic of human bone point
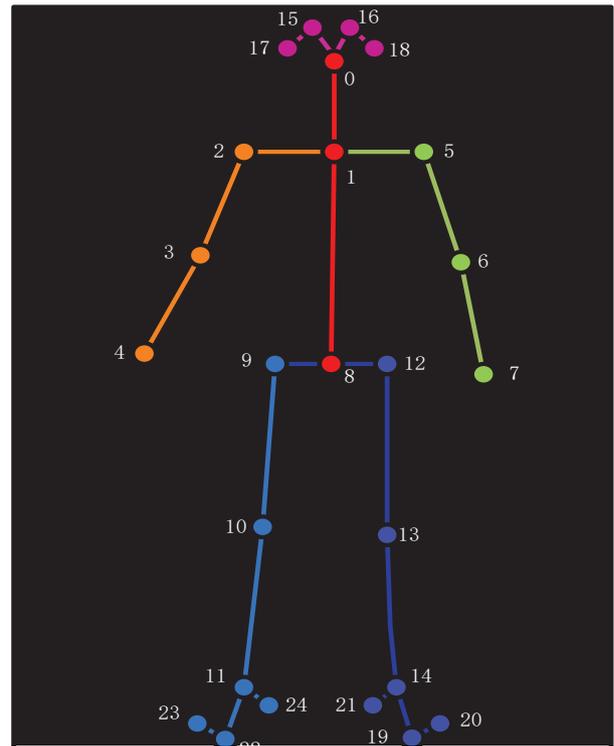
$$\tau = \varsigma / sum(\varsigma), \qquad (6)$$

where x refers to the input feature, y refers to the output feature, d refers to the total number of centralized bone points, δ refers to all-1 vector, τ refers to the feature vector of human body bone points, ς is sparse row vector, and sum is the summation function.

In Figure 1, there are 24 bone points. Except that the joint element value corresponding to the center point is expressed as 1, the element values of other bone points are set to 0. Set the number of centralized center points to 2, which are bone points 1 and 8, respectively, then the expression formula of sparse row vector is

$$\varsigma = [1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0], \qquad (7)$$

The sparse row vectors are obtained according to the above expression, and the feature data are added, modified and synthesized to obtain the optimized bone point feature extraction results.

## 2.3. Construction of Improved Deep Reinforcement Learning Recognition Model

To solve the limitations of traditional action recognition networks, based on STGCN network and bone feature optimization extraction results, the shallow neural network to DSTG dense connection network structure is improved in this paper, and an improved deep reinforcement learning network is obtained [21]. On this basis, a figure skating jumping action recognition model is constructed,as shown in Figure 3.

Figure 3 describes the overall structure of the DSTG network, which is mainly composed of FCN-BN-ReLu, three DSTG layers, two transition modules, one pool layer and one full connection layer. The optimized bone point features are input into the DSTG

block structure, and improved deep reinforcement learning is carried out through FCN-BN-ReLu, transition module and multiple DSTG layer units. To reduce the learning error, a feature dimension reduction layer is set between adjacent DSTG blocks. For the output value of the previous DSTG layer, the dimension-reduced feature is mapped to the low dimensional space as the input feature of the next DSTG layer [11]. Finally, through the operation of pooling layer and full connection layer, the jumping action recognition results are obtained. Figure 2(b) describes the DSTG layer structure, which is relatively simple and mainly composed of Concat, Convs, ReLu, Dropout and BN. Through the cooperation of different module supports, feature dimension reduc-

Figure 3

DSTG network structure and DSTG layer structure

(a) DSTG network structure

(b) DSTG layer structure
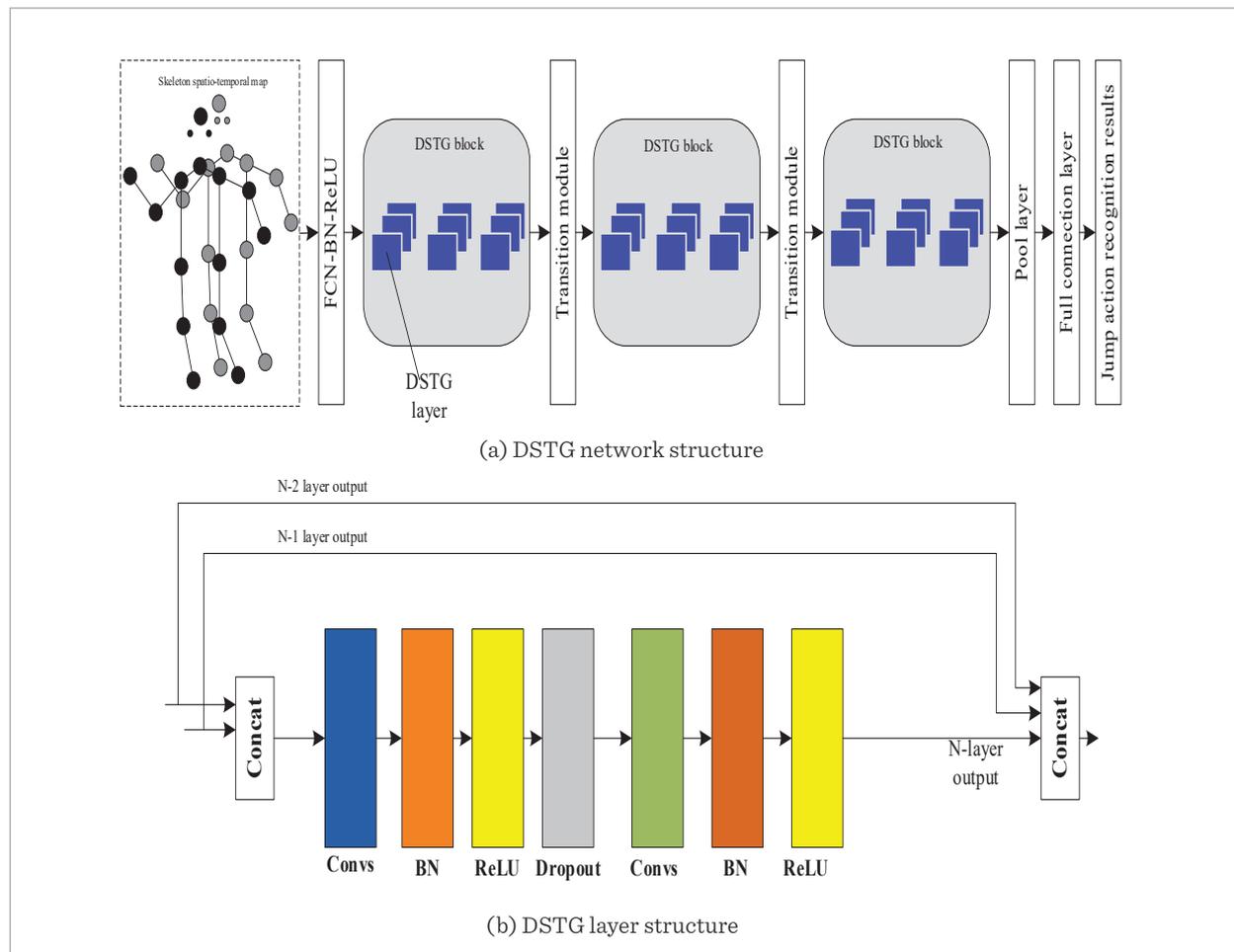
tion is realized, and the results are output to the next DSTG layer.

In the learning of motion recognition model, to improve the time convergence of the improved deep reinforcement learning network and reduce the gradient vanishing problem, the network parameters need to be initialized [18]. First, set the offset in all improved deep reinforcement learning network structures to 0. Then, the weight of the convolution layer in the network structure is initialized and analyzed to ensure that the final weight calculation results conform to the Gaussian distribution [5]. The minimum value of Gaussian distribution is 0, and the maximum value is calculated as follows:

$$\varphi = \sqrt{\frac{2}{(1+a^2) \times b}}, \tag{8}$$

where φ is the maximum value of Gaussian distribution, and b refers to the number of input parameters at the DSTG layer. Then, set the input parameters of the ReLu function to ensure that the weight variance remains unchanged during the forward propagation process.

Finally, the Adam algorithm is integrated into the improved deep reinforcement learning recognition model in this paper to optimize the random gradient descent. Then, the learning rate is attenuated by cosine annealing algorithm, and the cosine annealing calculation method is as follows:

$$\eta = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min})(1+\cos(\frac{\varpi_1}{\varpi})), \tag{9}$$

where η refers to the learning rate, η_{max} denotes the maximum learning rate, η_{min} refers to the minimum learning rate, ϖ is the total rounds of model training, ϖ_1 refers to the number of training rounds in the current cycle, and cos is cosine function.

Considering the periodicity of the learning rate change, Formula (9) is used to attenuate the learning rate three times to avoid that the learning rate is too small, resulting in the local optimization of the model identification results [7]. According to the above calculation results, the parameters of the recognition model are set, and then the figure skating videos are analyzed to get the jumping action recognition results.

2.4. Proposed Algorithm

To improve the recognition accuracy and efficiency, combined with the improved deep reinforcement learning recognition model, a jumping action recognition algorithm of figure skating videos is designed this paper, as shown in Figure 4.

Figure 4

Process of jumping action recognition algorithm for figure skating video
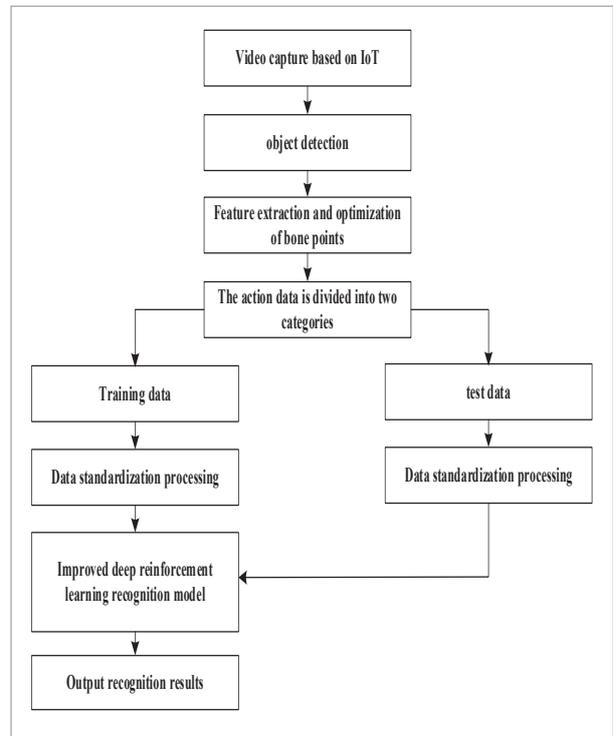
In this paper, the IoT technology is mainly to collect the figure skating video and detect the targets in figure skating videos, so as to obtain the human body bone point features through the feature extraction network, and carry out centralized processing to complete the optimization of the extraction results. The jumping action data in figure skating video is divided into training data and test data, and they are standardized, respectively. The standardized processing results of training data are input into the improved deep reinforcement learning action recognition model, the standardized processing results of test data are input into the trained model, and the jumping action

recognition results are output from the figure skating videos.

In figure skiing video, different time scales of action sequence division will cause differences in feature extraction results of spatial flow and motion flow, and affect the final jumping action recognition results. Based on the weighted fusion algorithm [3], the confidence fusion scheme is established, the recognition result with the highest confidence is selected, and the high-precision jumping action recognition result is output.

Under the function of the weighted fusion algorithm, the maximum probability and the second maximum probability of the current action samples belonging to the jumping action in figure skating videos are calculated, and the average probability error is calculated as the basis for calculating the confidence of the action recognition results. The expression Equation of confidence is

$$\gamma_j(i) = (1-\mu)(p_j^{\max}(i) - p_j^{sub\max}(i)) + \mu(p_j^{\max}(i) - \frac{1}{n}\sum_{i=1}^{n} p_j(i))$$ (10)

where i is the figure skating action sample to be recognized, j represents the category of jumping action, γ refers to the confidence degree, μ refers to the parameter of confidence degree, $p^{\max}$ refers to the maximum probability that the sample belongs to the jumping action, $p^{sub\max}$ is the second largest probability that the sample belongs to the jumping action, and n refers to the total amount of samples.

Considering that in jumping action recognition from figure skating videos, the time series are divided according to different time scales, the score vector and confidence of the original frame sequence and the score vector and confidence of the optical flow image sequence are required for each time scale. After weighted fusion calculation, the final score vector of jumping action recognition result is obtained.

$$\iota(i) = (\chi_1(i)W_1(i) + \chi_2(i)W_2(i)),$$ (11)

where ι refers to the output score vector of action recognition, χ_1 is the action recognition score vector of original frame sequence, χ_2 refers to action recognition score vector of optical flow image sequence, W_1

refers to the confidence degree of action recognition result of original frame sequence, and W_2 refers to the confidence degree of action recognition results of optical flow image sequences.

Based on Formula (11), for the figure skating video sample to be recognized, the finally recognized category label is expressed as

$$\upsilon(i) = \arg\max_j p_j^{\max} \cdot \iota(i),$$ (12)

where j refers to properties of action category labels, and $p_j^{\max}$ refers to parameter search function of maximum score. After the implementation of the confidence fusion scheme, the final jumping action recognition result from figure skating videos is obtained.

3. Experimental Analysis and Results

3.1. Data Set

Figure Skating data set: For World Figure Skating Championships in recent three years, collect the video of the single skating, divide all the videos into 1200 figure skating video clips, and set the video resolution to 720 × 360, and the frame rate of each video is 30 frames /s. To reflect the performance of the algorithm proposed in this paper, the data set includes three basic actions: jumping, rotating and stepping. In addition, according to the technical requirements of figure skating, the Figure Skating data set is divided into 10 categories of technical actions. An example of a dataset sample is shown in Figure 5.

The corresponding basic technical action categories of the 10 types of actions shown in Figure 5 are shown in Table 1.

MSR Action 3D data set: Use Kinect and other depth cameras to shoot athlete action videos in the figure skating training field. While the figure skating video is collected, the human 3D bone motion data can be obtained as the basic data for action recognition. The resolution of figure skating video contained in Giallo dataset is 720 × 360. All the actions are jumping, rotating and stepping actions performed by 10 athletes. To strengthen the authenticity of the experiment, each athlete should repeat the same action for 1-3 times.

Figure 5

Sample of Figure Skating data set
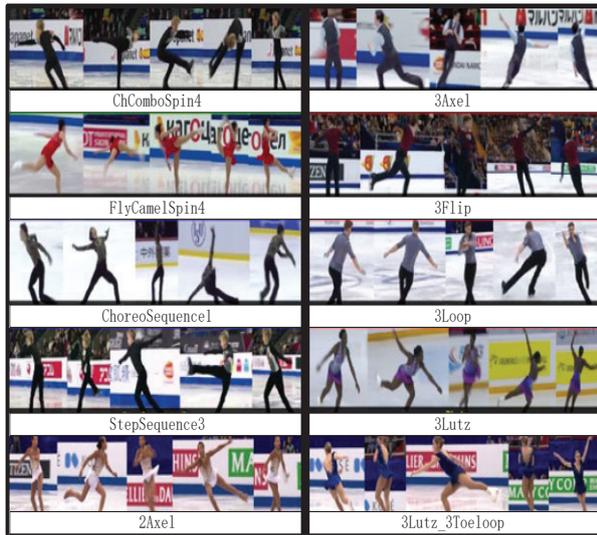
Table 1

Action category name and basic action classification

Category name	Basic technical action
2Axel	jump
3Axel	jump
3Flip	jump
3Loop	jump
3Lutz	jump
3Lutz_3ToeLoop	jump
ChComboSpin4	rotate
FlyCamelSpin4	rotate
ChoreoSequence1	pace
StepSequence3	pace

UCF 101 data set: This dataset is a commonly used benchmark dataset in the field of motion recognition, including various types of human action videos, all of which are sourced from YouTube video websites. In this experiment, figure skating jumping action is added to the data set and mixed with other action video samples for action recognition. Set the video resolution to 320 × 240 with Avi format stores all video samples.

HMDB51 data set: In movie clips and online videos, collect figure skiing videos and other human action videos, and get a total of 5000 videos. Similarly, set the resolution of the video segment to 320 × 240, uniformly saved as Avi format is used as the basic data for subsequent action recognition experiments.

3.2. Evaluation Metrics

1. In Figure Skating set, the algorithm proposed in this paper is applied to action recognition. The training data batch of the improved deep reinforcement learning network model is set to 32, and the learning rate and the number of extracted frames are 0.05 and 350 frames, respectively. In the video sample set, 80% of the figure skiing videos are selected for training, and 20% of the samples are selected for test. In addition to the recognition algorithm proposed in the paper, for Figure Skating data set, the algorithms in BRMFSJ [10], POSL [1], AHDM [16], ARMHJ [22] and HARM [13] are used for jumping action recognition, and then the true positive rate is calculated according to the action recognition results.

$$S = \frac{T}{T + F}, \tag{13}$$

where S refers to true positive rate, T refers to the number of samples identified as jumping actions in the test set, F refers to the total number of unrecognized jumping action samples. Generally, the larger the true positive rate is, the more accurate the recognition result is.

2. Aiming at MSR Action 3D dataset, select 80% of the samples to train the improved deep reinforcement learning network, and select 85% samples for training. After the training, for the remaining 15% of the action samples to test, the algorithm proposed in this paper and the algorithm proposed by literary are used to identify the jumping action. According to the recognition results of jumping actions and the collection of real jumping actions. To calculate IoU.

$$I = \frac{\overline{Y} \cap Y}{\overline{Y} \cup Y}, \tag{14}$$

where I refers to IoU, $\overline{Y}$ refers to the jumping action sets recognized, Y refers to true jumping action set, $\cap$ refers to intersection symbol, and $\cup$ refers to union symbol.

3 To facilitate the experimental analysis, the MSR Action 3D is enhanced to expand the figure skating video samples. 80% of the samples were randomly selected to train the improved deep reinforcement learning network, and 20% of the samples were selected for testing. Using the algorithm proposed in this paper, the recognition results of jumping motion are obtained. At the same time, five other recognition algorithms are applied to analyze the data set for action recognition. Calculate the F-measure value according to the recognition results of different algorithms.

$$F1 = \frac{2T}{2T + F + P}, \tag{14}$$

where $F1$ refers to the F-measure value, and P refers to the number of samples that are mistaken as jumping action.

4 According to the content proposed in this paper, 80% of the samples from UCF 101 data set are selected to train the improved deep reinforcement learning network, and 20% of the samples are selected for testing. Process the figure sketching video in the UCF 101 dataset, extract the features of human bone points, and obtain the feature extraction results shown in Figure 6.

The features are input to improve the deep reinforcement learning network model, and the jump action recognition results are obtained. At the same time, the algorithms in BRMFSJ [10], POSL [1], AHDM [16], ARMHJ [22] and HARM [13] are applied to perform action recognition on UCF 101 dataset. According to the recognition accuracy and recall of different recognition algorithms, the PR curve image is drawn, and then the area under the

Figure 6
Results of bone point feature extraction

| (a) Input Image | (b) Bipartite Matching | (c) Parsing Results |

curve is calculated by 11 point interpolation method to obtain the AP value.

5 Select 80% samples from HMDB51 data set to train the improved deep reinforcement learning network, and select 20% samples to test. The Receiver Operating Characteristic (ROC) curve is obtained based on the action recognition results. The abscissa of the curve is the false positive rate and the ordinate is the true positive rate. The AUC value is calculated according to the area under the ROC curve. Generally, the value range of AUC value is between 0 and 1. The closer to 1 indicates that the proposed algorithm has less recognition error.

3.3. Results and Discussion

The comparison results of the true positive rates of different algorithms are shown in Table 2.

According to the data in Table 2, we can see that with the increase of the number of experiments, the true positive rate of different algorithms shows an upward trend. When the number of experiments reaches 60, the true positive rate of different algorithms reaches the maximum value. The true positive rate of algorithms in the paper is 0.98, which is 0.23, 0.26, 0.24, 0.32, 0.29 higher than the algorithms in BRMFSJ [10], POSL [1], AHDM [16], ARMHJ [22] and HARM [13], respectively, It shows that compared with the experimental comparison algorithm, the true positive rate of the algorithm in this paper is higher, which proves that the recognition result of this algorithm is more accurate.

The comparison results of IoU of different algorithms are shown in Table 3.

According to the data in Table 3, we can see that with the increase of the number of experiments, the intersection and merging ratio of different algorithms shows a fluctuating trend. The maximum IoU of the algorithm in the paper is 0.97, which is 0.26, 0.27, 0.24, 0.29, 0.3 higher than that of the algorithms in BRMFSJ [10], POSL [1], AHDM [16], ARMHJ [22] and HARM [13], respectively. This shows that the IoU of the algorithm in the paper is higher than that of the algorithm in experiment comparison, it shows that the action recognition result of this algorithm is closer to the reality.

The comparison results of F-measure values of different algorithms are shown in Figure 7.

By analyzing the data in Figure 7, we can see that the
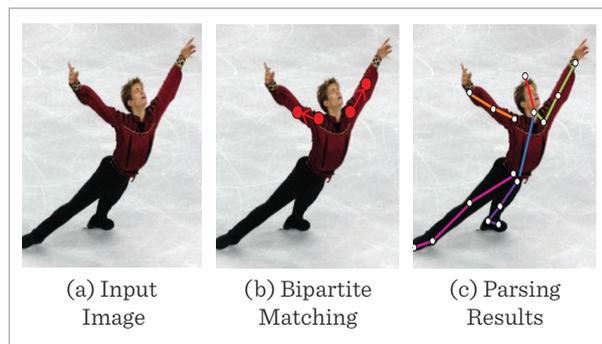
Table 2

Comparison results of true positive rate

Number of experiments	BRMFSJ [10]	POSL [1]	AHDM [16]	ARMHJ [22]	HARM [13]	Proposed
10	0.61	0.51	0.53	0.46	0.51	0.89
20	0.63	0.59	0.59	0.47	0.53	0.93
30	0.64	0.63	0.63	0.52	0.58	0.95
40	0.66	0.64	0.64	0.56	0.61	0.96
50	0.72	0.67	0.68	0.63	0.63	0.97
60	0.75	0.72	0.74	0.66	0.69	0.98

Table 3

Comparison results of IoU

Number of experiments	BRMFSJ [10]	POSL [1]	AHDM [16]	ARMHJ [22]	HARM [13]	Proposed
10	0.67	0.69	0.71	0.61	0.63	0.94
20	0.71	0.61	0.69	0.65	0.62	0.94
30	0.69	0.70	0.63	0.67	0.59	0.95
40	0.61	0.59	0.73	0.63	0.63	0.95
50	0.62	0.51	0.72	0.66	0.65	0.96
60	0.63	0.52	0.69	0.68	0.67	0.97

Table 4

Comparison results of AP values

| Number of experiments	BRMFSJ [10]	POSL [1]	AHDM [16]	ARMHJ [22]	HARM [13]	Proposed
10	0.71	0.66	0.61	0.61	0.51	0.92
20	0.73	0.67	0.64	0.62	0.53	0.91
30	0.78	0.68	0.65	0.60	0.52	0.93
40	0.72	0.71	0.63	0.58	0.53	0.92
50	0.69	0.72	0.61	0.61	0.58	0.91
60	0.72	0.70	0.62	0.62	0.57	0.93

F-measure values of different algorithms show a fluctuating trend with the increase of the number of experiments. The maximum F-measure value of the algorithm in the paper is 0.98, which is 0.18, 0.23, 0.28, 0.33 and 0.34 higher than the algorithms in BRMFSJ [10], POSL [1], AHDM [16], ARMHJ [22] and HARM [13], respectively. It shows that compared with the experimental comparison algorithm, the F-measure value of the algorithm in the paper is higher, indicating that the action recognition effect of the algorithm is better.

The comparison results of AP values of different algorithms are shown in Table 4.

According to Table 4, we can see that the maximum AP value of the algorithm in this paper is 0.93, which is 0.15, 0.21, 0.28, 0.31 and 0.37 higher than the algorithms in BRMFSJ [10], POSL [1], AHDM [16], ARMHJ [22] and HARM [13], respectively. It shows that compared with the experimental comparison algorithm, the AP value of the algorithm in this paper is higher, indicating that the action recognition rate of this algorithm is higher.

Figure7

Comparison results of F-measure values different algorithms

Figure 8

AUC value comparison results
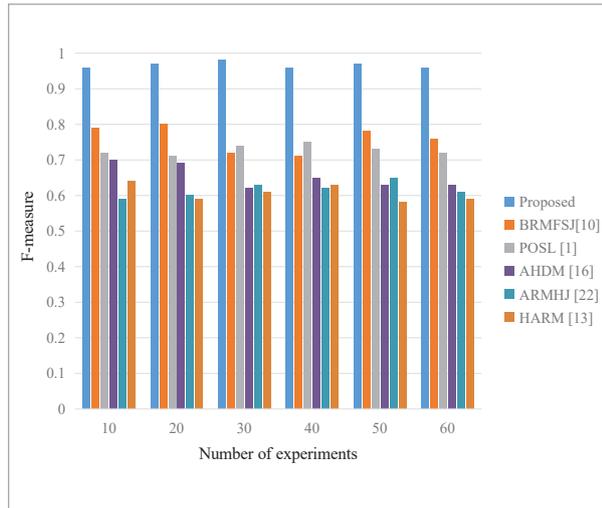
The comparison results of AUC values of different algorithms are shown in Figure 8. We can see that the maximum AUC of the algorithm in this paper is 0.96, which is 0.19, 0.17, 0.21, 0.3 and 0.3 higher than the algorithms in BRMFSJ [10], POSL [1], AHDM [16], ARMHJ [22] and HARM [13], respectively. It shows that compared with the experimental comparison algorithm, the AUC value of the algorithm in this paper is higher, indicating that the action recognition effect of this algorithm is better.
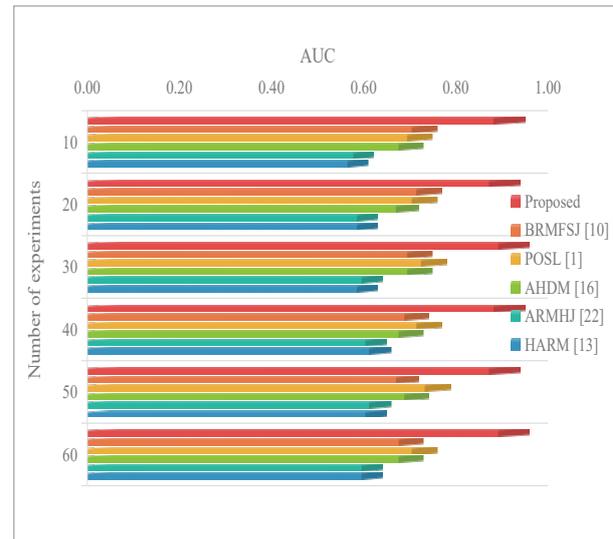
To sum up, the maximum true positive rate of the algorithm in this paper is 0.98, the maximum IoU is 0.97, the maximum F-measure is 0.98, the maximum AP is 0.93, and the maximum AUC is 0.96, which shows that the recognition accuracy of this algorithm is higher, and the recognition effect is better. It can effectively improve the accuracy of the recognition results of figure skating video jumping action, and can be widely used in the field of figure skating action recognition to ensure the training quality.

4. Conclusions

Taking the jumping action in figure skating video as the research object and with the rapid development of the IoT, a recognition algorithm based on improved deep reinforcement learning network is proposed.

The experimental results show that the AP value and AUC value of the proposed algorithm reach more than 0.9, the true positive rate and IoU are 0.98 and 0.97, respectively, and the F-measure value of the algorithm is 0.98, which is 18%, 23%, 28%, 33% and 34% higher than the other five recognition algorithms. In the experimental results, the proposed algorithm can more accurately recognize jumping actions from the video, and can be used as an important auxiliary tool for figure skating teaching and training. Although the proposed algorithm of using improved deep reinforcement learning for jumping action recognition in figure skating video in IoT is promising, there are still some potential shortcomings and limitations that need to be addressed. For instance, the algorithm's performance in recognizing certain types of jumping actions or under specific conditions might be suboptimal, and it does not consider other factors that could affect athletes' performance or training outcomes. Therefore, future research should explore the possible challenges and drawbacks of this algorithm and develop strategies to improve its effectiveness and applicability.

Acknowledgement

References

1. Delisle-Houde, P., Chiarlitti, N, Reid., R, Andersen, E. Predicting On-Ice Skating Using Laboratory- and Field-Based Assessments in Collegiate Ice Hockey Players. International Journal of Sports Physiology and Performance, 2019, 1184-1189. https://doi.org/10.1123/ijspp.2018-0708

2. Feng, D., Wu, Z., Zhang, J., Ren, T. Multi-Scale Spatial Temporal Graph Neural Network for Skeleton-Based Action Recognition. IEEE Access, 2021, 9(1), 58256-58265. https://doi.org/10.1109/ACCESS.2021.3073107

3. Feng, L., Zhao, Y., Zhao, W., Tang, J. A Comparative Review of Graph Convolutional Networks for Human Skeleton-based Action Recognition. Artificial Intelligence Review, 2021, 55(5), 4275-4305. https://doi.org/10.1007/s10462-021-10107-y

4. Feng, H., Zhao, X., Zhang, X. Automatic Arrangement of Sports Dance Movement Based on Deep Learning. Computational Intelligence and Neuroscience, 2022, 2022, 1-14. https://doi.org/10.1155/2022/9722558

5. Gullu, H H., Sirin, D S., Yldz, D E. Analysis of Double Gaussian Distribution on Barrier Inhomogeneity in An Au/n-4H SiC Schottky Diode. Journal of Electronic Materials, 2021, 50, 7044-7056. https://doi.org/10.1007/s11664-021-09254-3

6. Ha, M H., Chen, T C. Deep Neural Networks Using Capsule Networks and Skeleton-Based Attentions for Action Recognition. IEEE Access, 2021, 9(1), 6164-6178. https://doi.org/10.1109/ACCESS.2020.3048741

7. Islam, M S., Bakhat, K., Khan, R., Iqbal, M., Islam, M., Ye, Z. Action Recognition Using Interrelationships of 3D Joints and Frames Based on Angle Sine Relation and Distance Features Using Interrelationships. Applied Intelligence, 2021, (51), 6001-6013. https://doi.org/10.1007/s10489-020-02176-3

8. Inzamam M.N, Mudassar R, Jamal H.S, Wang, S., Tariq, U., Muhammad, A. HAREDNet: A Deep Learning Based Architecture for Autonomous Video Surveillance by Recognizing Human Actions. Computers and Electrical Engineering, 2022, 99, 1-17. https://doi.org/10.1016/j.compeleceng.2022.107805

9. Jiang, F., Chen, X. An Action Recognition Algorithm for Sprinters Using Machine Learning. Mobile Information Systems, 2021, 2021(5), 1-10. https://doi.org/10.1155/2021/9919992

10. Mazurkiewicz A. Biomechanics of Rotational Movement in Off-Ice Figure Skating Jumps: Applications to Training. Polish Journal of Sport and Tourism, 2021, 28(2), 3-7. https://doi.org/10.2478/pjst-2021-0007

11. Mohammadzade, H., Hosseini, S., Rezaei-Dastjerdehei, M., Tabejamaat, M. Dynamic Time Warping-Based Features With Class-Specific Joint Importance Maps for Action Recognition Using Kinect Depth Sensor, IEEE Sensors Journal, 2021, 21(7), 9300-9313. https://doi.org/10.1109/JSEN.2021.3051497

12. Mallick, T., Das, P., Majumdar, A. Posture and Sequence Recognition for Bharatanatyam Dance Performances Using Machine Learning Approaches. Journal of Visual Communication and Image Representation, 2022, 87, 103548. https://doi.org/10.1016/j.jvcir.2022.103548

13. Nguyen N, Pham, M., Pharm, C., Tachibana K. Human Action Recognition Method Based on Conformal Geometric Algebra and Recurrent Neural Network. Information and Control Systems, 2020, 5(1), 2-11. https://doi.org/10.31799/1684-8853-2020-5-2-11

14. Shehzad, F., Attique, M., Asfand, E., Sharif, M., Alhaisoni, M., Tariq, U., Majumdar, A. Two-Stream Deep Learning Architecture-Based Human Action Recognition. Computers, Materials & Continua, 2023, 74(3), 5931-5949. https://doi.org/10.32604/cmc.2023.028743

15. Voronin, V., Zhdanova, M., Semenishchev., Zelenskii. A., Cen, Y., Agaian, S. Action Recognition for the Robotics and Manufacturing Automation Using 3-D Binary Micro-Block Difference. The International Journal of Advanced Manufacturing Technology, 2021, (4), 1-12. https://doi.org/10.1007/s00170-021-07613-2

16. Wang, S., Tong, S. Analysis of High-level Dance Movements Under Deep Learning and Internet of Things. J Supercomput, 2022, 78, 14294-14316. https://doi.org/10.1007/s11227-022-04454-y

17. Wang, W. Research on the Architecture of Digital Song and Dance Costume Design System Based on Intelligent Deep Learning Algorithm. Journal of Sensors, 2022, 2022, 1-10. https://doi.org/10.1155/2022/1545334

18. Xu, J., Ai, B. Deep Reinforcement Learning for Handover-Aware MPTCP Congestion Control in Space-Ground Integrated Network of Railways, IEEE Wireless Communications, 2021, 28(6), 200-207, 50(12), 7044-7056. https://doi.org/10.1109/MWC.001.2100116

19. Xie, J., Miao, Q., Liu, R., Xin, W., Tang, L., Zhong, S., Gao, X. Attention Adjacency Matrix Based Graph Convolutional Networks for Skeleton-Based Action Recognition. Neurocomputing, 2021, 440, 230-239. https://doi.org/10.1016/j.neucom.2021.02.001

20. Xia, Z H., Xing, J.M., Li, X F. Gesture Tracking and Recognition Algorithm for Dynamic Human Motion Using Multimodal Deep Learning, Security and Communication Networks, 2022, (2022),1-11. https://doi.org/10.1155/2022/4387337

21. Yin, H., Wan, J., Zhang, S J., Xu, Z. ADSCN: Adaptive Dense Skip Connection Network for Railway Infrastructure Displacement Monitoring Images Super-resolution. Multimedia Tools and Applications, 2021, 80(13),6105-6120. https://doi.org/10.1007/s11042-020-10009-1

22. Zhai, L., Duan, H., Chen, D. An Adaptive Recognition Method for Take-off Action Images of Back-style High Jump Based on Feature Extraction. Future Generation Computer Systems, 2021, 126(1), 65-69. https://doi.org/10.1016/j.future.2021.07.032

23. Zheng, Y, Liu, Z., Lu, T.,Wang, L. Dynamic Sampling Networks for Efficient Action Recognition in Videos. IEEE Transactions on Image Processing, 2020, 29, 7970-7983. https://doi.org/10.1109/TIP.2020.3007826