Research Article

An Action Recognition Technology for Badminton Players Using Deep Learning

Jiatong Liu[1] **and Bo Liang** [2]

[1]*Physical Education Department, Zhejiang Wanli University, Ningbo 315100, China*
[2]*Physical Education Department, Shandong Technology and Business University, Yantai 264005, China*

Correspondence should be addressed to Bo Liang; 201513285@sdtbu.edu.cn

With the fast development of sports in recent years, the number of people participating in various sports has increased day by day. Because of the advantages of fewer constraints on the field and ease of learning, badminton became one of the most popular sports among them. Numerous works have been done specifically for the recognition of action of badminton players to improve and popularize it, but the traditional badminton player's badminton action recognition algorithm employs the method of manually constructing a topology map to model the action sequence contained in multiple video frames. Besides, it learns each video frame in a targeted manner to reflect data changes, which is prone to high computational cost, low network generalization, and catastrophic forgetting. In response to the above problems, this paper proposes a deep learning-based action recognition technology for badminton players, which re-encodes the human hitting action sequence data with multirelational characteristics into relational triples, and learns by decoupling based on long short-term memory network. At the same time, this paper designs and completes a set of badminton action recognition schemes based on acceleration and angular velocity signals. Experimental results show that the proposed method achieves 63%, 84%, and 92%, respectively, recognition accuracy on multiple benchmark datasets, which improves the accuracy of human hitting action recognition. As a result, the evaluation will be useful in future work to improve the structure of current deep learning models for higher results in badminton action recognition.

1. Introduction

Badminton is an Olympic event and among the most famous racket sports in the world. Badminton is characterized by the efficiency of the game with no timeframe, ambiguity, higher explosive criteria, speed, and a wide range of sports. It is an essential item in modern people's daily lives in sporting events, recreational opportunities, and entertainment. To some extent, these factors have aided the fast expansion of badminton programs. These days, data collection during playing badminton and subsequent data analysis to complete action recognition have emerged as a hotspot of intelligence in the sports industries. Motion recognition-based products and services are also becoming increasingly popular. It is currently the mainstream of the sports to identify the data of a user's movement, state, exercise volume, and other data through the data created by a user playing badminton game, and to provide the user with some healthy exercise advice based on this information [1–8]. As a result, badminton requires extensive study directions, which improves the game.

In the field of action recognition during playing badminton game, there are two major study directions. The first is based on inertial sensors, which recognize actions. For example, some researchers collect user action data using three-axis acceleration sensors and then utilize a hidden Markov model for time series modeling to realize gesture recognition based on the data collected [9, 10]. Alternatively, the principal component analysis (PCA) method is employed for human motion detection based on the data received by the three-axis acceleration sensor. The other option is to utilize video monitoring and image analysis to

detect and identify suspicious activity. Action features are extracted from the skeleton motion trajectory and combined with other data to form an action recognition model to realize action recognition. In comparison to video-based action recognition methods, the human skeleton sequence data contains more accurate joint position information, making the action recognition approach based on human skeleton data more robust when dealing with perspective change, body proportion, and motion speed [11–16].

Traditionally, on the badminton court, camera equipment is used to watch the movement of athletes to capture the posture of the human body, which is then evaluated using image processing software. Because this technology necessitates the use of high-quality camera equipment and extensive image processing, the action detection cycle is lengthy, and the cost is significant, making it unsuitable for usage in regular badminton games. When it comes to badminton players, the HSAE-based technique surpasses the video-based approach when it comes to action detection. Handcrafted feature-based approaches, which are manually designed features, capture the dynamic relationship between joints and thereby capture the dynamic interaction between joints [17, 18]. Recently, deep learning approach has become milestone in action recognition. Deep learning-based methods are used to represent the joints of various sections of the human body end-to-end using deep models such as recurrent neural networks. The latter is composed primarily of three frameworks: Neuronal networks are used to model joint sequences that are structured under particular principles in sequence-based approaches. Image-based approaches employ convolutional neural networks to model human skeleton sequences encoded as pseudoimages, whereas graph-based methods use graph-based methods to model human skeleton sequences encoded as pseudoimages. It uses the joint points of the human body as vertices and the natural connections between the joint points as edges to construct a topology map for the joint data. At the same time, the high-dimensional features of the skeleton data are used to construct the human skeleton data, which is then used to construct the human skeleton data using a graph convolutional neural network [19–23].

It is necessary to manually generate a topology map for joint data before supplementing it with high-dimensional features using the graph-based method. An action in a real-world scenario, on the other hand, is typically constituted of many video frames. Manually creating a topology network, which connects joints, bones, and joint points and to reflect dependencies between disconnected joint nodes, is a time-consuming process that requires skill [24, 25]. To recognize each video frame in an action, the model must be trained individually for that frame, and the parameters must be changed regularly to reflect data changes. This will result in a higher computational cost and a significant potential for catastrophic forgetfulness [18, 26, 27].

This study proposes a deep learning-based badminton player action recognition technique to address the aforementioned issues. In addition, the LSTM in our model dynamically generates the topological map of the human skeleton. Moreover, the HSAE is implemented using a mod-

eled model that is based on the concept of continuous learning. The following are the most significant contributions made by this work:

(1) In this paper, we proposed to use a dynamic deep learning-based action recognition technology to recognize badminton players. In addition, we implemented an assessment mechanism during the training phase, which guarantees that the network is capable of reliably identifying the learned action categories. The skeleton topology map is dynamically generated by decoupling, and an evaluation mechanism is introduced during the training process. Identify and resolve the catastrophic forgetting problem that occurred during training

(2) Using the concept of continuous learning, the network learns the decoupling features of the sequence frame entities according to action categories, and a partial update strategy to construct a dynamically updated human skeleton from badminton player skeleton data with multirelational characteristics. By decreasing the network's complexities, the topology map aids in its generalization

(3) Finally, our experimental findings obtained using publicly available datasets demonstrate that the method described in this research can significantly increase the accuracy of action recognition tasks performed by humans

The remaining sections of this paper are arranged as follows: Badminton hitting theories and techniques are explained in Section 2; our proposed methodology for recognizing badminton player's action using deep learning is explained in Section 3; the experimental work and the results obtained from our experimental work are explained in Section 4; and our work's conclusions are presented in Section 5.

2. Badminton Hitting Theory and Techniques

2.1. Hitting a High Ball with the Forehand. Hitting a high ball with a forehand in the backcourt is a fundamental action in the game of high ball tennis. To properly estimate the direction of the incoming ball and the landing place, we must first teach pupils how to do so accurately. Students should elevate the position over the right shoulder such that the left shoulder faces the net and the left foot is in contact with the ground. Stand with your right foot in front of you, and your entire weight should be supported by your right foot. Students should maintain a natural rise of their left arm and flexed left hand. Alternately use a right-handed racket while keeping their upper and lower arms in a natural bend. Elevate the racket above the right shoulder, and gaze at the badminton court with both eyes open. To hit the ball properly, the upper arm should be led back, the joint raised, and the joint raised to the top of the shoulder. The racket should be led up to the back of the head in order to ensure that the wrist is naturally stretched and the heart of

the first is facing upward, as in the preparation action. Ensure that the rear foot is pushing the ground, allow the waist and abdomen to coordinate force, and utilize the shoulder as an axis. Let the big arm drive the forearm, causing it to travel forward swiftly, and strike the ball at the highest point of the straight arm.

Despite the recognition of the shot, the inertial force causes the arm holding the shot to swing forward and downward, closing the shot to the front of the body. Meanwhile, the left foot is pushed further back, while the right foot is pushed more forward. It is also necessary to transfer the body weight from the right foot to the front foot during this process. A forehand strike on a high-bouncing ball may result in a different shot. Taking off and hitting high and long balls is a strategy for attempting to hit the highest hit ball and win the initiative, which necessitates greater physical power and footwork. Beginners are often interested in learning how to strike high and long balls without having to hop.

Following the shot's impact, the inertial action causes the arm holding the shot to swing forward and downward, closing the shot to the front of the body. This has a direct impact on the hitting force. If you are striking the ball, the first stage is to use the shoulder as the axis to swing the arm, and then, switch to using the elbow as the axis. This will alter the effect of the weak arm, resulting in inappropriate force. The ball is pushed out by force. After hitting the ball, the racket is not tilted downwards, and the racket is tilted downwards to the front of the body. Instead, it is swung backward, which affects the force of the arm after striking the ball. When the racket surface is not properly positioned when striking the ball, the racket is struck obliquely or tangentially, lowering the quality of the stroke. When they hit the ball, they take quick steps back, causing them to lose sight of the hitting point because their bodies are facing forward and they are moving back with parallel steps.

2.2. High Ball Was Hit over the Head. Hit the badminton with your forehand above the left shoulder and in the middle of the court, and hit the opponent's bottom line with a high ball with your forehand on the left rear court. Chinese athletes were the first to develop this type of technique when it came to hitting action. When compared to the backhand striking technique, this approach has a higher sense of initiative and a greater sense of aggression. The high and long shots are both the same distance from the target. It is important to note that the high and long shots are taken from the top of the left shoulder, and the ball is launched out to the left-back corner. Using your left shoulder to hit the ball, bend back slightly to the left, and place your left shoulder towards the net. When hitting the ball, the big arm should drive the forearm over the top of the head, and swing forward swiftly from the upper left. The big arm should also force the ball into the ground. It is important to remember to use the explosive power of the wrist when hitting the ball, to fully push the ground, and to use the strength of the abdomen while hitting the ball. Immediately after the shot, place your left foot on the ground and push back, then move your right foot forward and move your weight back to your right

foot to finish the shot. Students need coordination, footwork, and balance for this amazing style. As students master the fundamentals of mobility, they should increase the frequency with which they exercise so that they can use them flexibly in practice, competition, and on the backcourt. Strengthen your offensive aggression in the area.

The faults that students frequently make during training are comparable to those that occur when forehand a high and long ball with the forehand. They do not maintain their sideways position when hitting the ball, do not employ their arm strength, and do not apply the entire force of their body, resulting in uncoordinated hitting and insufficient force. Because the player hits the ball at the incorrect spot, the hitting point is either too far ahead or too far behind the intended location. A low hitting point is produced by having the elbow strike the ball at an incorrect angle or too wide. This has an impact on the overall hitting quality. At the same time, the surface of the racket is improper, and the racket is tilted or sliced, which will have an impact on the quality of the shot taken. Simultaneously, the position of the center of gravity when striking the ball is too far back, which will affect the combination's high and long shots, as well as the time required to prepare for the following catch, which will be delayed.

2.3. Basic Techniques in Badminton. To become a professional badminton player, players must understand some fundamental skills like basic capital or beginning capital. Figure 1 illustrates the basic techniques of badminton.

2.3.1. Holding a Racket. The simplest technique of playing badminton is how to capture a racket. It is a significant factor influencing the performance of the most manufactured punches. Anyone who wants to learn badminton must first learn this basic skill before learning the next one. The catching technique for this racket contains two types of methods, as described below:

  (i) Technique of forehand: With this technique, the method of holding the racket can be done with the right or left hand. During this technique, hold the racket as if shaking hands with other people, with a space between the thumb and index finger formed like a letter, and the other fingers gripping the racket

  (ii) Technique of backhand: This technique is identical to the forehand technique; besides, the backhand technique can be used with either the right or left arm to grasp the racket. The backhand technique similarly holds the racket to the forehand technique, with the exception that the thumb and index finger are near. Position of the racket head. The thumb's position differs slightly from the other four fingers, and four fingers hold the racket

2.3.2. Hit. The four-hand technique is the most common hit technique in the game of badminton or badminton. In the four-hand technique, the blow pulls the body back to the front of the racket, where the palm faces the shuttle cocktail. Grasp the racket in the forehand position, and place the
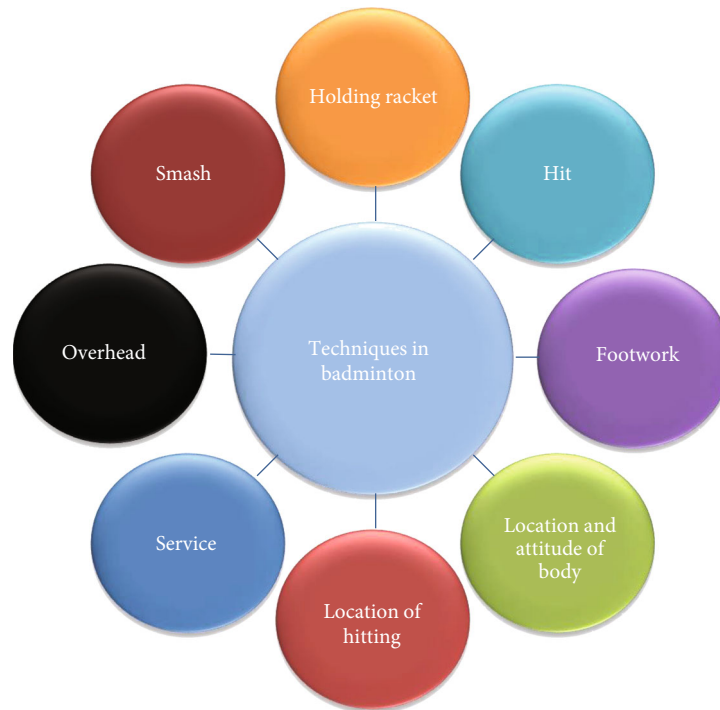
FIGURE 1: Techniques for badminton.

right and left legs during this technique. Adjust to the right with your body. Make a point of hitting the shuttle cocktail as tight as you can.

2.3.3. Footwork. Agility is the one basic ability or basic talent that must be acquired in badminton. A badminton player must be flexible in their mobility, not only with their hands but also with their other limbs, such as their foot motions. This is because harmonization of hands, body, and legs is essential to do and produces quality while also delivering devastating blows. We require good cooperation among the parts of the body not only to carry out an attack but also to sustain the opponent's attack.

2.3.4. Location and Attitude of Body. The principle of balancing is important to the basic badminton method of attitude and body posture. Not only is quickness vital in badminton, but so is equilibrium. These are the simple activities that are frequently overlooked by average folks. On the other hand, a professional badminton player understands the importance of balance in the game. Best badminton competitions at the native, national, and global levels provide real support for the importance of the stability factor in the game of badminton. A badminton person's balancing act is often poor, causing them to be less than optimum level or even failed to generate quality punches.

2.3.5. Location of Hitting. The quality of shots in badminton is highly influenced not only by hitting but also by the body position during hitting. Attempt to keep the body on one side of the net in the meantime. The left foot is in front of the right foot, the body is behind the shuttle cocktail, the

right shoulder is significantly forced back, and the right shoulder and right foot positions should be altered when throwing blows.

2.3.6. Service. The service technique in badminton is to move the shuttle cock to the opposing player's right, left, front, or back. The duty of the shuttlecock is something to avoid while serving, although it is fine in front of the opponent. This is akin to suicide because the opponent will be able to quickly restore service by eliminating our motions.

2.3.7. Technique of Overhead. In badminton, overhead occurs when the shuttlecock's falling position leads to the rear of our standing body posture. Furthermore, the racket is caught using the forehand method, and the overhead shuttle cocktail can be hit as well as thrown.

2.3.8. Technique of Smash. Smash is an attacking movement technique in badminton that aims to hit the opponent's movements. The best smash will be made if the high jump is made because we can get a perfect smash position in this position. This sabotage is carried out with force, and the shuttle cocktail is launched into the opponent's territory. Because it is done with full force, the shuttle cock moves very quickly on the smash.

3. Methodology

3.1. Basic Mathematical Work. Mathematics is essential to all of our areas of expertise. It facilitates the quantitative measurement of principles and constructs, the formulation of assumptions and postulates, the analysis of our

observational data, and the presentation of indisputable scientific findings.

3.1.1. Human Skeleton Sequence (HSS).

All the bones and joints of the body are part of the human skeletal system. Each bone is a complex living organism consisting of numerous cells, protein fibers, and minerals. The skeleton acts as a support, supporting and protecting the connective tissue that comprises the rest of the body. The skeletal system also acts as a connecting point to the muscles, allowing them to move in the joints. HSS is usually composed of consecutive human skeleton frames; each human skeleton frame is a set of 2D or 3D coordinates of a series of joint points and the confidence of the coordinates. Let $HSS = \{HS_1, HS_2, L, HS_s\}$, where S_s is the number of HSS frames, and HS_i is the ith human skeleton frame in HSS. HS_i consists of a series of joint points containing spatiotemporal information, namely, $HSS_i = \{H_1, H_2, \cdots, H_N\}$, where N is the number of joint points in the HSS frame. Define the ith joint point of the HSS frame H_i, where (x_i, y_i) is the position information of the joint point and the $score_i$ is the confidence information of the joint point position. The dimension of HSS is (T, N, C). Among them, T is the number of HSS frames, N is the number of joint points, and C is the dimension of the position vector (the dimension of the position vector of S_i is 3).

3.1.2. Space-Time Graph and Multirelationship Graph.

A space-time graph is a pictorial depiction of the badminton hitting technique theory of relativity's characteristics of time and space. It enables testing the significance of relating occurrences such as time travel and length contraction in the absence of mathematical equations. In order to facilitate the use of graph convolution to extract the spatial and temporal features of the HKS, the HKS is usually reconstructed into a spatiotemporal graph $G = (V_i, E_i)$, which contains the joint point set V_i and the edge set E_i. The joint point set $V = \{Vt/t = 1 \cdots T, i = 1, \cdots, N\}$, including all the joints in the human skeleton sequence.

In all frames, the edge set $E = \{E_S, E_F\}$, composed of people, the body skeleton connection E_S and the frame E_F are composed of two subsets. The first one is the intraframe HKS connection, which is given by $E_S = \{V_{ti}V_{tj} | (i, j) \in H\}$, where H is the naturally connected human joint set. While, the second subset is the interframe connection, which is given by $E_F = \{V_{ti}V_{(t+1)i}\}$, the motion trajectory of a specific joint in E_F can be represented by all of the edges of that joint over the course of time.

3.1.3. Relational Triples.

A graph structure with many-to-many relationships between nodes is called a multirelational graph, and a multirelational graph can be formalized as a set of relation triples $\{(u, r, v)\} \subseteq V \times E \times V$.

3.2. Overall Structure Design.

The sensor for the data collecting module is mounted on the bottom of the badminton racket and the wrist of the human body, respectively. A digital signal is created from an analog signal collected during movement by using analog-to-digital conversion, which is

then imported into the motion recognition system at a sampling frequency of 10 Hz. Signal preprocessing is completed after digital filtering and eigenvalue calculation, and the swing capture technique is utilized to extract the eigenvalues using active window postdetection after signal preprocessing has been completed. Figure 2 depicts a design diagram of the system.

In order to address the issues of insufficient generalization and catastrophic forgetting in the process of learning new data in traditional action recognition algorithms based on graph convolution, this paper incorporates the concept of continuous learning and proposes a deep learning-based algorithm based on a GCN network to address the issues of insufficient generalization and catastrophic forgetting. Badminton players can benefit from action recognition technology. A sequence of human skeleton triples is constructed from the multirelational characteristics of HKS itself. Decoupling is used to learn embedded features on multirelational data, and partial update techniques are used to dynamically reconstruct the relationship topology of HKS. When it comes to action recognition, a topological map with many relationships is input into a GCN network for processing. Figure 3 depicts the overall architecture of the algorithm, which is subdivided into four sections: data preprocessing, feature decoupling, dynamic topology map creation, and action category prediction.

It is mostly low-frequency signals that are used to measure the acceleration and angular velocity of badminton motion. The signal acquisition apparatus utilized in this work operates at a frequency of 10 hertz (hertz). As stated by the Nyquist sampling theorem, the greatest frequency of the collected signal is 5 hertz (hertz). During the processes of excitation, detection, and transmission, the noise will have an impact on the resulting signal. Furthermore, as a result of the jitter of the racket during the movement of the ball, interference noise will be introduced. As a result, the action signal must be filtered in order to remove the interference caused by high-frequency background noise. The bilinear change approach is utilized in this research to develop the Butterworth low-pass IIR digital filter, which is used to filter the input signal.

In actual sports, the racket is utilized in any direction independent of the ball's orientation. As a result, it is illogical to utilize the three-axis acceleration signal and the three-axis angular velocity signal as the eigenvalues of the action without first calculating the eigenvalues of the action. In order to minimize the interference caused by the action direction, it may be necessary to calculate the resultant acceleration signal and the resultant angular velocity signal. It is recommended to calculate the variance of the acceleration signal to be able to see different movements when swinging due to the high magnitude of the acceleration and angular velocity generated. So to summarize, the eigenvectors of this work are constructed by taking into account their variance, the angular velocity signal they produce, and their own acceleration signal as the eigenvalues of their respective eigenvectors.

Therefore, the three-axis acceleration signal collected at time t is recorded as a_{xt}, a_{yt}, a_{zt}, and the three-axis angular
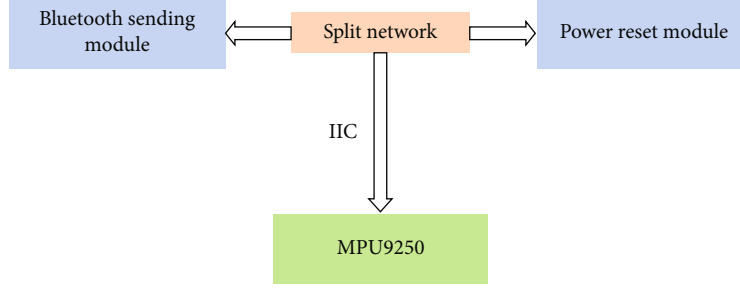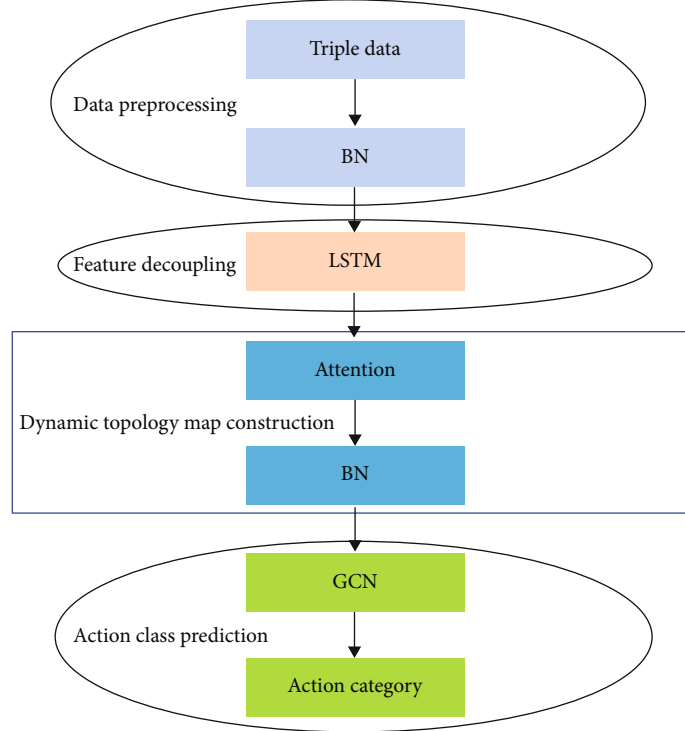
FIGURE 2: Action data acquisition module.

FIGURE 3: Structure diagram.

velocity is j_{xt}, j_{yt}, j_{zt}. Calculate the resultant acceleration A, acceleration variance D_t, and resultant angular velocity J_t at time t, respectively. These are given in

$$A_t = \sqrt{a_{xt}{}^2, a_{yt}{}^2, a_{zt}{}^2},$$
$$J_t = \sqrt{j_{xt}{}^2, j_{yt}{}^2, j_{zt}{}^2}, \tag{1}$$
$$D_t = \left(\left|a_{xt}\right| - A_t\right)^2 + \left(\left|a_{yt}\right| - A_t\right)^2 + \left(\left|a_{zt}\right| - A_t\right)^2.$$

Data preprocessing first normalizes the joint feature vectors in the dataset composed of multiple HKS HK_i, converts HKS in different positions to the same position to promote better convergence of the algorithm, and then reprocesses the dataset composed of multiple HKS HK_i. Specifically, according to the multirelational characteristics of the HKS data itself, the HKS relation triple sequence $RS = \{(HK_i, r_{ij}, HK_j)\}$ is defined, where HK_i is the set of joint points of the i-th frame in the HKS set, HK_j is the set of joint points in the jth frame in HKS, and r_{ij} is the action category represented by the HKS set.

When using feature decoupling, the dynamic topology map building module takes advantage of the concept of continuous learning and creates skeleton topology maps for various sorts of activities depending on the feature vectors acquired. To be more specific, the K-means algorithm is used to cluster the relational feature embedding vectors representing action categories that are produced by feature decoupling, and the dataset is separated into several training sets based on the clustering centers that are discovered. Following one-hot encoding, the attention mechanism and partial update technique are utilized to generate topology maps of different sorts of actions based on the sequence of action categories that have been encoded. A similar approach is used to evaluate model performance once each batch of data training has been completed, with action recognition accuracy serving as a measure to ensure that the model can still
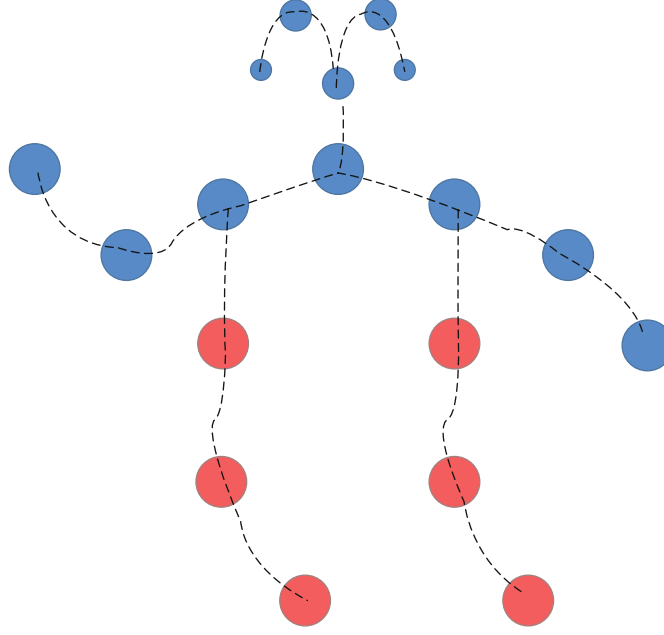
FIGURE 4: Diagram of feature.

produce better evaluation results on the learned dataset after learning new data. The importance of joint points is defined in

$$\alpha_r^i = \frac{\exp\left(a_r^i\right)}{\sum_{j=1}^N \exp(a_r^i)}. \tag{2}$$

When judging the distance, the method is given in

$$d(X, Y) = \sqrt{(X - Y)^T \sum{}^{-1}(X - Y)},$$
$$d(X, G) = \sqrt{(X - \mu)^T \sum{}^{-1}(X - \mu)}, \tag{3}$$
$$W_{ij} = d^2\left(X_k, G_j\right) - d^2\left(X_k, G_i\right).$$

At 30 frames per second, the Microsoft Kinect detector creates a depth image with a pixel size of 640 480. Human body shapes are provided in real time by the produced depth map action scenes. The 20 body joints are approximated from the body style by using the technique from [28], and the skeletal system is designed from the depth map series. The skeletal model, also recognized as the "stick model," includes 20 body joints with specified joint axes, as shown in Figure 4.

The activated loss function is can be calculated by using

$$L_1 = - \sum_{\mathrm{HS}_i \in N(S_i)} \frac{1}{|c|} y(c) \cdot \log\left(\sigma(w\mathrm{VS}_i)\right). \tag{4}$$

Equation (5) represents the loss function of the regularization term L_2.

$$L_2 = \sum_{\mathrm{HS}_i \in N(S_i)} \left(1 - \sum_i^n \alpha_r^i\right). \tag{5}$$

Finally, the overall loss can be calculated by using

$$L = L_1 + \beta L_2. \tag{6}$$

4. Experiment Results

The approach in this study is tested using the HKSR dataset skeleton and the NTU dataset to ensure that it performs as expected. This dataset contains all of the HSK key point information in the movie that was detected by DeepMind through the associated estimation software and stored as part of the skeleton dataset. There are 400 categories of actions in the dataset, which contains approximately 300,000 videos in total. Each frame of skeleton data contains 18 human body joint point locations and confidence scores, and each frame of skeleton data contains 18 human body joint point locations and confidence scores. The NTU dataset classifies activities into 60 categories, with a total of approximately 70,000 movies. These videos include RGB videos with a resolution of 19201080, depth map sequences with a resolution of 512424, infrared videos, and 3D videos with 25 human body joints, among other types of videos. The location's 3D skeleton data has been uploaded.

All algorithms in the experiment are chosen based on the parameters that are the most efficient. The bidirectional LSTM coding layer and the softmax classification layer are the two most important layers in the model presented in this research. It is used to extract features from the data, and then, the softmax classification layer is used to forecast the action category, which is then used to evaluate the model's effectiveness. The Adam optimizer is used to train the model
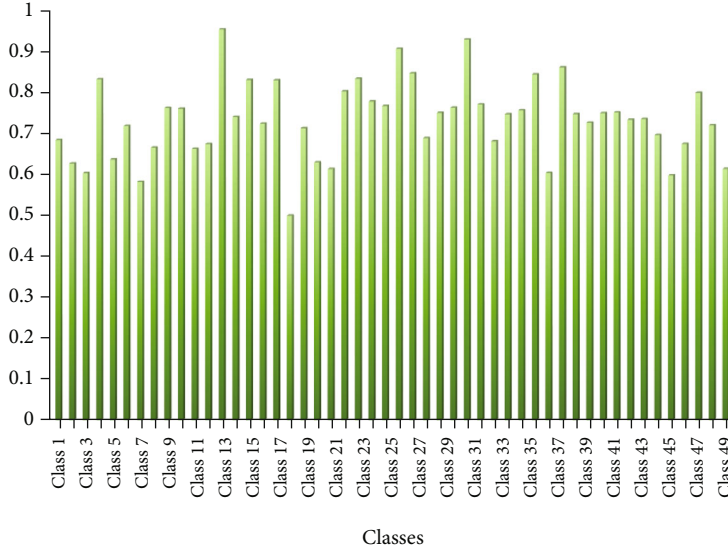
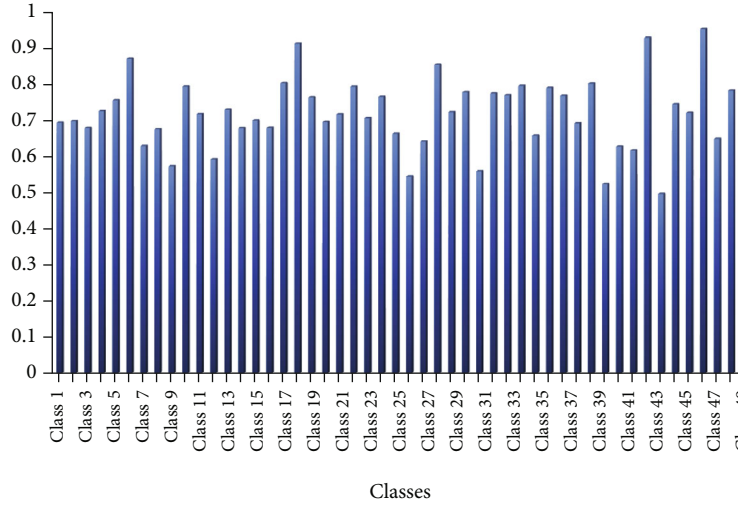FIGURE 5: Accuracy of our algorithm on skeleton.

FIGURE 6: Accuracy of our algorithm on NTU.

in this experiment, with an $r = 0.001$ accuracy. Initialize the bidirectional LSTM encoding layer's parameter settings with the default values. Max length = 300, input size = 75, and hidden size = 200 are all valid values. Initialize the parameters of the softmax classification layer: the input vector dimension is the same as the output vector dimension of the bidirectional LSTM, the number of categories is the same as the number of action categories in the dataset, the bias term and random deactivation are not used, and the hyperparameter of the regularization term in the loss function is the same as the loss function hyper parameter. Set to 0.1; other comparison algorithms make use of the best settings available for their respective parameters when doing comparisons.

On the skeleton dataset and on the NTU dataset, the dynamic topology graph building block is utilized to dynamically construct multirelationship graphs, which are pre-

sented in this section. The softmax layer is used to evaluate the accuracy of the output action category after the model has been input into each category dataset. The Top-1 accuracy of the output action category is evaluated using the softmax layer after the model has been input into each category dataset. Top-1 accuracy rates for each action category are shown in Figures 5 and 6 for the two datasets, respectively.

The performance of deep learning-based action recognition approaches for badminton players is evaluated in this part from the perspectives of overall performance and average performance, as well as the performance of individual players. During the evaluation of overall performance, all datasets are supplied into the network at the same time, and the recognition accuracy throughout the full test set is determined. When evaluating average performance, the test set is fed into the network by category, and evaluation is carried out during the training process, after which the average
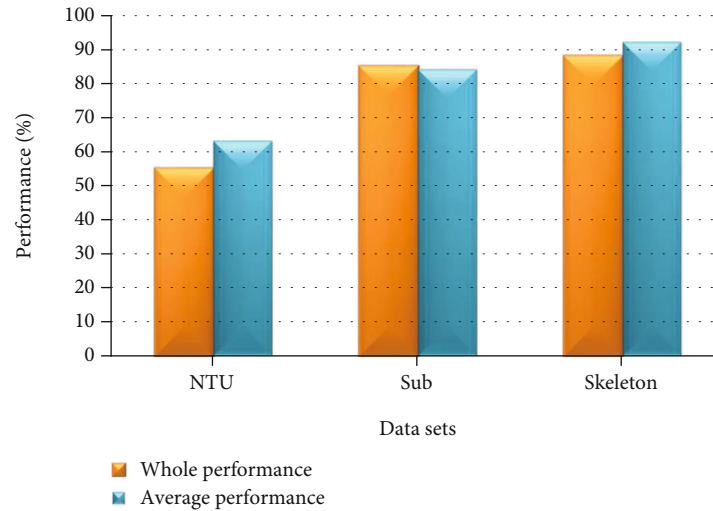
FIGURE 7: Comparison between whole and average performance.

of all test set evaluation metrics is obtained. Because the average performance may be used to emphasize the performance of the algorithm when dealing with the catastrophic forgetting problem, it is the primary metric used to evaluate the model's performance.

On the three datasets, it can be seen in Figure 7 that the average performance of the deep learning-based badminton player action recognition technology is better than the overall performance, which fully demonstrates that the dynamic topology map construction strategy proposed in this paper can input the datasets into the network by category. Specific action categories are learned more intensively by the network than other categories of actions, thereby decreasing the impact of other categories of actions on the current recognition task, solving the problem of catastrophic forgetting during the process of updating network parameters, and increasing the generalization of the network.

5. Conclusions

Sport performance analysis is an important branch of sports practice. To recognize the action of badminton players and analyze its performance, this paper proposes a deep learning-based action recognition system for badminton players. Feature embedding is learned through decoupling, which is based on the idea of continuous learning. A partial update mechanism is used to dynamically construct a human skeleton topology map when dealing with triples of new skeleton relations, and a GNN-based skeleton action recognition algorithm is used to achieve action recognition. The problems of high computational cost, low network generalization, and catastrophic forgetting are effectively addressed in this paper. On multiple benchmark datasets, experimental results show that the proposed method achieves 63%, 84%, and 92%, respectively, recognition accuracy, improving the accuracy of human hitting action recognition.

Data Availability

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

Conflicts of Interest

Declares that he has no conflict of interest.

References

[1] N. A. Rahmad and M. A. As'ari, "The new convolutional neural network (CNN) local feature extractor for automated badminton action recognition on vision based data," *Journal of Physics: Conference Series*, vol. 1529, no. 2, article 022021, 2020.

[2] Y. Wang, W. Fang, J. Ma, X. Li, and A. Zhong, "Automatic badminton action recognition using cnn with adaptive feature extraction on sensor data," in *International Conference on Intelligent Computing*, pp. 131–143, Springer, Cham, 2019.

[3] N. A. binti Rahmad, N. A. J. binti Sufri, M. A. bin As'ari, and A. binti Azaman, "Recognition of badminton action using convolutional neural network," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 7, no. 4, pp. 750–756, 2020.

[4] N. A. Rahmad, M. A. As'ari, M. F. Ibrahim, N. A. Sufri, and K. Rangasamy, "Vision based automated badminton action recognition using the new local convolutional neural network extractor," in *International Conference on Movement, Health and Exercise*, pp. 290–298, Springer, Singapore, 2020.

[5] J. Chen, R. D. J. Samuel, and P. Poovendran, "LSTM with bio inspired algorithm for action recognition in sports videos," *Image and Vision Computing*, vol. 112, article 104214, 2021.

[6] M. F. Tsai and S. H. Huang, "Enhancing accuracy of human action recognition system using skeleton point correction method," *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 7439–7459, 2022.

[7] A. Ghosh, S. Singh, and C. V. Jawahar, "Towards structured analysis of broadcast badminton videos," in *2018 IEEE Winter*

Conference on Applications of Computer Vision (WACV), pp. 296–304, Lake Tahoe, NV, USA, 2018.

[8] H. Kim, M. Jain, J. T. Lee, S. Yun, and F. Porikli, "Efficient action recognition via dynamic knowledge propagation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 13719–13728, Ximen, 2021.

[9] L. Fang and M. Sun, "Motion recognition technology of badminton players in sports video images," *Future Generation Computer Systems*, vol. 124, pp. 381–389, 2021.

[10] N. Gao, "Construction and simulation of athlete's wrong action recognition model in sports training based on embedded wireless communication and computer vision," *Wireless Communications and Mobile Computing*, vol. 2021, 11 pages, 2021.

[11] Z. Chu and M. Li, "Image recognition of badminton swing motion based on single inertial sensor," *Journal of Sensors*, vol. 2021, 12 pages, 2021.

[12] E. P. Ijjina, "Action recognition in sports videos using stacked auto encoder and HOG3D features," in *Proceedings of the Third International Conference on Computational Intelligence and Informatics*, pp. 849–856, Singapore, 2020.

[13] Y. Wang, G. Pan, J. Ma, X. Li, and A. Zhong, "Label similarity based graph network for badminton activity recognition," in *International Conference on Intelligent Computing*, pp. 557–567, Springer, Cham, 2021.

[14] S. Liwei, "Research on classification and recognition of badminton batting action based on machine learning," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 5, pp. 6241–6252, 2019.

[15] A. Raj, P. Consul, and S. K. Pal, "Fast neural accumulator (NAC) based badminton video action classification," in *Proceedings of SAI Intelligent Systems Conference*, pp. 452–467, Cham, 2021.

[16] S. Ramasinghe, "Poster: advanced feature based deep learning for intelligent human activity recognition: an approach using scene context and composition of sub events," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services Companion*, p. 75, Bolon, 2016.

[17] B. Martinez, D. Modolo, Y. Xiong, and J. Tighe, "Action recognition with spatial-temporal discriminative filter banks," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5482–5491, Beijing, 2019.

[18] Z. Tu, W. Xie, J. Dauwels, B. Li, and J. Yuan, "Semantic cues enhanced multimodality multistream CNN for action recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 5, pp. 1423–1437, 2019.

[19] L. Song, G. Yu, J. Yuan, and Z. Liu, "Human pose estimation and its application to action recognition: a survey," *Journal of Visual Communication and Image Representation*, vol. 76, article 103055, 2021.

[20] M. R. Tseng, A. Gupta, C. K. Tang, and Y. W. Tai, "HAA4D: few-shot human atomic action recognition via 3D spatio-temporal skeletal alignment," 2022, https://arxiv.org/abs/2202.07308.

[21] P. K. Singh, S. Kundu, T. Adhikary, R. Sarkar, and D. Bhattacharjee, "Progress of human action recognition research in the last ten years: a comprehensive survey," *Archives of Computational Methods in Engineering*, vol. 28, pp. 1–41, 2021.

[22] H. Lee, Y. S. Kim, M. Kim, and Y. Lee, "Low-cost network scheduling of 3D-CNN processing for embedded action recognition," *IEEE Access*, vol. 9, pp. 83901–83912, 2021.

[23] S. Cao and R. Nevatia, "Exploring deep learning based solutions in fine grained activity recognition in the wild," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 384–389, Cancun, 2016.

[24] B. S. Chawky, A. S. Elons, A. Ali, and H. A. Shedeed, "A study of action recognition problems: dataset and architectures perspectives," in *Advances in Soft Computing and Machine Learning in Image Processing*, pp. 409–442, Springer, Cham, 2018.

[25] M. F. Hashmi, B. K. K. Ashish, and A. G. Keskar, "GAIT analysis: 3D pose estimation and prediction in defence applications using pattern recognition," *Twelfth International Conference on Machine Vision (ICMV 2019). International Society for Optics and Photonics*, vol. 11433, p. 114330S, 2020.

[26] E. E. Cust, A. J. Sweeting, K. Ball, and S. Robertson, "Machine and deep learning for sport-specific movement recognition: a systematic review of model development and performance," *Journal of Sports Sciences*, vol. 37, no. 5, pp. 568–600, 2019.

[27] S. Jain, A. Rustagi, S. Saurav, R. Saini, and S. Singh, "Three-dimensional CNN-inspired deep learning architecture for yoga pose recognition in the real-world environment," *Neural Computing and Applications*, vol. 33, no. 12, pp. 6427–6441, 2021.

[28] J. Shotton, A. W. Fitzgibbon, M. Cook et al., "Real-time human pose recognition in parts from single depth images," in *Computer Vision and Pattern Recognition*, vol. 1297-1304, Colorado Springs, CO, USA, 2011.