IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 2 November 2023, accepted 18 December 2023, date of publication 21 December 2023,
date of current version 27 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3345636

RESEARCH ARTICLE

The Analysis of Technical Characteristics of Badminton for Sports With Neurorobotics Under Machine Learning

FAN WU AND HUAN CHEN
Nanchang University College of Science and Technology, Gongqingcheng 330029, China
Corresponding author: Huan Chen (6203119030@email.ncu.edu.cn)

ABSTRACT With the increase in public attention to sports and health, more people are keen to use badminton for fitness, but it is difficult to find professionals for guidance. In the new era, technology is constantly updated, and the artificial badminton sports intelligent computing platform has encountered difficulties in software and hardware engineering technology from data collection to computing programming. Firstly, this paper studies the imaging mechanism and calibration method of the binocular camera and obtains the relationship between the parameters and the position of the lens in the camera. Secondly, a convolutional computing network model Region Proposal Network (RP-ResNet) is established, and the first-stage feature image data collection is performed on the input image. The second step uses an improved pyramid pooling network to pool feature images at different resolutions. Besides, Squeeze-and-Excitation Network (SENet) enhances channel attention, thereby improving the measurement accuracy of small target objects in the network. Finally, a dynamic identification algorithm of badminton hitting based on the sliding window is given, and a real-time identification system of badminton dynamics is implemented on this basis. A statistical system of badminton technical characteristics is established through these badminton dynamic identification methods. According to the test results, the algorithm model system designed here can identify common hitting actions in real-time. The improved Hidden Markov Model (HMM) improves the comprehensive recognition accuracy by 1.25% and shortens the recognition time by 0.07s compared with the traditional HMM. This model will provide an intelligent data analysis platform for badminton, which can be promoted and applied to professional players and coaches. The technical index parameters suitable for the regular development of badminton sports are extracted from the accumulated big data analysis to establish a professional product for intelligent data analysis and auxiliary training of badminton sports.

INDEX TERMS Binocular vision, convolutional neural network, action recognition, badminton action feature statistics, neurorobotics, machine learning.

I. INTRODUCTION

Artificial intelligence technology is a comprehensive emerging discipline developed by the interpenetration of computer technology, cybernetics, bioinformatics, and other disciplines [1], [2]. It is a science and technology that understands the inner mechanism of human intelligence and conducts practical research and development on machines. Machine learning is also the main development core application area of the new generation of artificial intelligence technology research and development, and its application has penetrated all the main branch application areas of the new generation of artificial intelligence technology [3], [4]. In recent years, with the development of artificial intelligence technology and machine learning, a new generation of artificial intelligence technology has continued to innovate and iterate. The application of artificial intelligence technology in various industries is rapidly expanding, forming a new round of development of the "artificial intelligence +" industry revolutionary upsurge [5]. A new generation of artificial intelligence has

The associate editor coordinating the review of this manuscript and approving it for publication was Petros Nicopolitidis.

also emerged in traditional sports applications. Many products for data analysis and smart sports management related to the popular badminton sport have also been born, and it has initially shown the charm of applying a new generation of artificial intelligence technology to traditional sports [6]. However, the current badminton training methods are mainly one-to-one or one-to-many manual teaching methods.

There is a lack of targeted training methods for players to give full play to their advantages and overcome their shortcomings. In addition, the skills of badminton's auxiliary teaching function are the most demanding [7], [8]. The badminton robot has its visual system, so it is necessary to further develop the capture, recognition, and analysis system of the badminton player's swing action on the existing visual system and expand the function of the badminton automation robot [9]. In the development of a system for capturing, identifying, and analyzing badminton players' swing actions, image recognition skills for badminton swing actions are particularly important. It is crucial to identify how to capture and analyze badminton swing dynamics through the identification method of badminton racket [10].

Badminton has become a popular sport because of its moderate exercise intensity, easy-to-use rules, social interaction and indoor and outdoor characteristics. This wide popularity makes the analysis of badminton technology very important. Technical analysis can help improve the level of athletes, personalize training, optimize competition strategies and prevent sports injuries. Therefore, for badminton players and coaches, technical analysis is a crucial tool to improve the competition level, enhance the competition experience, and promote the healthy development of badminton. This paper expounds on the vision system of the neural robot, introduces binocular vision, and uses the principle of triangulation to obtain the positioning relationship between the camera and the target object during shooting to determine the exact position of the target. Then, the mechanism of the badminton robot's capture of the athlete's actions is illustrated, and the features of the convolutional neural network are used for in-depth cognition. Also, an improved pyramid pooling structure is used to improve the efficiency of network features and the measurement accuracy of small objects. The improved Hidden Markov Model (HMM) model is built, and a real-time recognition system for badminton actions is implemented. The system can effectively realize the capture and analysis of badminton players' actions and the recognition function of hitting actions. The main goal of badminton action recognition research is to capture, analyze and identify badminton players' various hitting trends and postures in the competition in real time. The motivation of the research stems from two main problems: Firstly, with people's increasing concern about sports health, more and more people choose badminton as a fitness activity, but it is often difficult for them to find professionals to guide them. Secondly, with the development of technology in the new era, the application of artificial intelligence technology in badminton is expected to provide better training and technical support. Therefore, the primary goal of this paper is to solve the above problems, and provide a technical basis for the intelligent data analysis and the development of auxiliary training products in the field of badminton, and promote the intelligent development of badminton. The realization of this goal will help to provide professional technical guidance and personalized training, and provide better technical support for badminton lovers and professional players. This paper presents a unique and innovative solution for real-time recognition and analysis of badminton movements by combining the vision system of the neural robot, binocular vision, convolutional neural network, and an improved pyramid pooling structure. This method leverages advanced technologies and integrates them, making their application in badminton possible. Furthermore, by establishing an improved HMM, this paper delves into the deep integration of modern technology and intelligence, providing non-experts with professional-level technical guidance and personalized training. Moreover, it offers unprecedented intelligent support for the field of sports technology. This comprehensive innovative approach holds the potential to drive technological and intelligent advancements in the world of badminton, offering higher-level support to athletes and badminton enthusiasts.

II. ACTION DETECTION AND RECOGNITION OF BADMINTON ROBOT

Action detection and recognition in badminton robots is a technology aimed at enabling robots to monitor real-time badminton and player movements on the court. It involves identifying and understanding different actions, such as the trajectory of the shuttlecock and the posture of players, to assist the robot in taking appropriate actions and responses, such as hitting the shuttlecock or adjusting its position. This requires the use of computer vision and machine learning techniques to process video data, extract relevant features, and train models for accurate action recognition and response. The task of action detection and recognition in badminton robots is complex, involving computer vision and machine learning technologies. To achieve this, a large amount of badminton motion video data needs to be collected and annotated. Then, training and validation are performed through feature extraction and the use of deep learning models like the convolutional neural network or recurrent neural network. Once the training is complete, the model is deployed on the robot to monitor shuttlecocks and players on the court in real-time, identifying actions and taking appropriate actions. This task requires highly accurate computer vision and machine learning technologies to meet the requirements of various application scenarios and constant model refinement and optimization to enhance accuracy and robustness.

A. BADMINTON ROBOT AND BINOCULAR VISION SYSTEM

Badminton robots can cooperate with people to complete real-time competitions [11], [12]. The visual control system of the badminton robot is mainly composed of two parts. The first part is the support, which is used to fix the camera and screen. The second part is the camera, which uses a binocular stereo vision control system to obtain the three-dimensional position of the badminton on the playing field [13], [14].

The further realization of key technologies such as machine vision and autonomous driving is inseparable from the external signals collected by depth cameras. Like ordinary cameras, this camera can also directly shoot the target object to measure the actual distance of the object [15], [16]. Binocular vision is a common way to extract external signals in machine vision. Both cameras are on the same baseline. Extract object signals at different positions based on the camera coordinate system. The difference in image signals is obtained by dual positioning, and the orientation signal and three-dimensional information of the target object are calculated [17]. Fig. 1 shows a typical depth camera.

FIGURE 1. Typical depth camera.

The binocular in-depth camera used in this paper has low power consumption and relatively perfect technology. The basic principle of image imaging of the binocular depth camera is to match the main feature points of the image captured by the camera with the projected light spot. Then, the triangulation method is used to get the positional relationship between the photographic camera and the target object to determine the correct position of the target [18]. Assuming that the front and rear positions of the camera are at the same X-Z axis position, the position of the target p in the camera coordinate system is p (x_0, y_0, z_0). The projected positions of the target p on the two camera positions are $p_1(x_1, y_1, 0)$ and $p_r(x_r, y_r, 0)$ in turn. Since the two cameras are on the same horizontal line, y_l is equal to y_r. The position of the projected point is $p_1(x_1, y_1, 0)$ and $p_r(x_r, y_r, 0)$. f is the focal length of the lens. The center distance of the two cameras is A, which is also the center baseline distance of the binocular camera. Therefore, both cameras form a closed triangle with the target object. The following equations can be acquired according to the triangular relationship of the triangle.

$$X_1 = f\frac{x_c}{y_c} \tag{1}$$

$$X_r = f\frac{(x_c - A)}{z_c} \tag{2}$$

$$y = f\frac{y_c}{z_c} \tag{3}$$

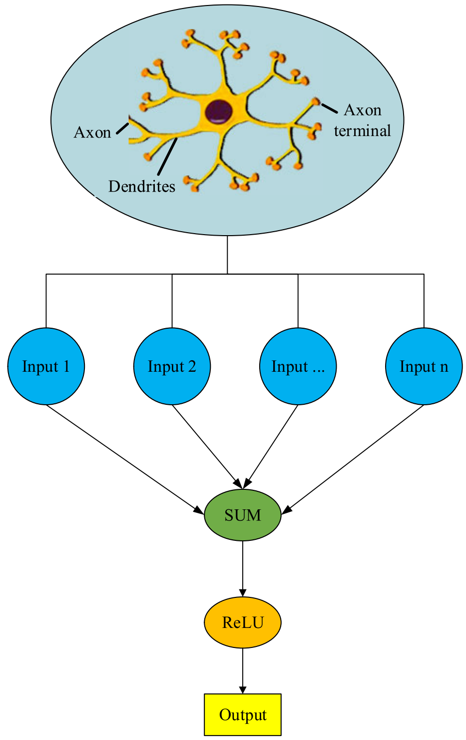


FIGURE 2. Structure of animal neurons and human simulated neurons.

Then, the relative position of the target position p in the camera coordinate system can be obtained.

$$x_0 = A\frac{x_l}{x_l - x_r} \tag{4}$$

$$y_0 = A\frac{y}{x_l - x_r} \tag{5}$$

$$z_0 = A\frac{f}{x_l - x_r} \tag{6}$$

According to the above basic principles, the three-dimensional position of the target point can be estimated to determine the depth information after the feature points are successfully matched.

B. PERCEPTION MODEL

Animal neurons and human simulated neurons are demonstrated in Fig. 2.

The left side of Fig. 2 displays the structure of animal neurons [19], [20]. Neurons are the basic structures for animals to capture and transmit information. On the right side of Fig. 2 is the researchers' virtual neuron architecture [21]. The network structure is formed by the end-to-end connection of neurons layer by layer. The input information is emulated as $x = (x_1, x_2, \ldots, x_n)^T$. The corresponding weight of the input signal is expressed as $\omega = (\omega_1, \omega_2, \ldots, \omega_n)^T$. After the given information is operated in the neuron, the calculation is completed by the activation function f. The computer outputs the connection y.

$$y = f\left(\sum\nolimits_{i=1}^{n} \omega_i x_i - \theta\right) \tag{7}$$

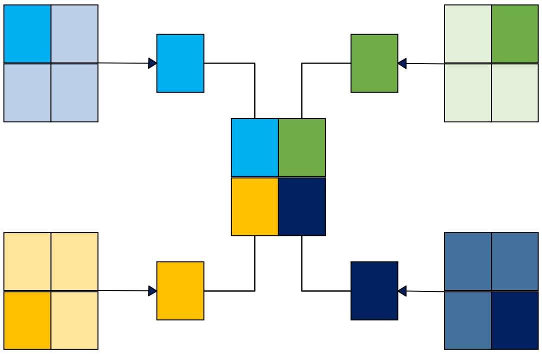


FIGURE 3. Schematic diagram of convolutional layer.

C. BASIC STRUCTURE OF THE CONVOLUTIONAL NEURAL NETWORK

The basic structural units of convolutional neural networks are convolutional layers, pooling layers, and fully connected layers [22], [23].

1) CONVOLUTIONAL LAYERS

The convolution algorithm layer is the basic structure of the convolutional neural network. The feature map can be generated through the convolution algorithm, and it can be expressed as (8) by mathematical methods.

$$x_j^l = f\left(\sum\nolimits_{i \in M_j} x_j^{l-1} * k_{ij}^l + b_j^l\right) \tag{8}$$

In (8), x_j^l represents the jth feature image of the lth layer. $*$ represents the convolution operator. M_j represents the set of all feature images to be convolved. k_{ij}^l represents the convolution kernel parameter in the output result of the layer. b_j^l represents the bias term of the jth feature image of the lth layer. f represents activation function. The schematic diagram of the convolutional layer is shown in Fig. 3.

2) POOLING LAYERS

Pooling layer can also be called down-sampling layer, which includes many different types of nonlinear functions. In practical application, the two common methods are maximum pooling and average pooling [24], [25].

The maximum pooling strategy refers to taking the maximum value of the filter as the output result and using the convolution kernel of N×N for pooling operation [26]. Although the maximum pooling can greatly reduce the scale of data information, large variation can also cause information loss or excess. The average pooling strategy means that the output result is the average value of all elements in the filter range [27]. The average pooling can be used when all the information in the feature has a certain contribution. For example, when it is located in a deep place in the network, because the width of the feature map is small, there are more semantic information implied.

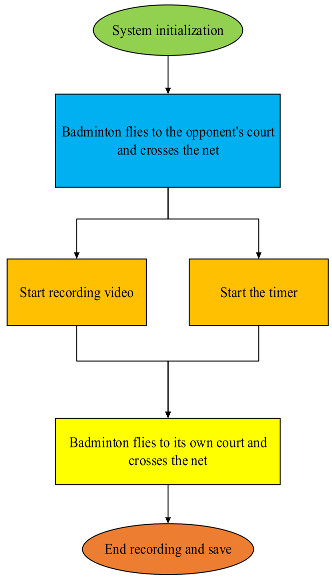


FIGURE 4. Frame diagram of video capture algorithm.

3) FULLY CONNECTED LAYER AREA

Each neuron in the fully-connected layer is connected to the neuron in the previous network, and the difference data in the convolutional and pooling layers can be utilized. In addition, data signals at each level are further processed through activation functions, thereby reducing parameters and improving network stability [28], [29].

D. REGION PROPOSAL NETWORK (RP-RESNET) MODEL

The two stages of region proposal network (RP-ResNet) model are as follows:

The first stage of the test is influenced by the regional suggestion network technology, and shared the feature images of the middle convolution layer (the 30th layer) of ResNet-50. The shared convolution layer and the output of regional suggestion network are injected into the Region of Interest (ROI) together, and the re-statistics of ROI on the region of interest are cancelled [30].

In the second stage, after the 30th floor of the network, Squeeze-and-Exclusion Network (SENet) is added to improve the pyramid pooling architecture. Senet templates are added after each residual detection module to strengthen the communication between channels. The higher-level signals are integrated through the improved spatial pyramid pooling architecture to improve the efficiency of network characteristics and improve the measurement accuracy of small objects to further improve the measurement accuracy [31].

E. CAPTURE OF SWING MOTION

According to the classification of the badminton action process, the badminton swing dynamics can be divided into three important stages, which are the preparation stage, the hitting stage, and the return stage. For the badminton robot, the capture is started when the shuttlecock flies towards the opponent's court and crosses the net. The capture is terminated when the shuttlecock flies back to the field and crosses the net. Fig. 4 shows the frame diagram of the modified badminton swing capture program based on the video capture program.

As Fig. 4 shows, a timer is used when the action starts. The purpose is to solve the problem that the shuttlecock cannot fly back to the area where the robot is located, so the dynamic shooting completion information cannot be obtained, thereby suppressing the shooting of the video. After the capture time threshold T_{max} is set, if the capture time is greater than the threshold, it is confirmed that the shot is wrong. Meanwhile, the collection of the video segment is terminated, and the video content of this segment is discarded. The optional value of the threshold can be determined according to the maximum time required for the swing action. After the total time of each video segment in the badminton swing dynamic data set is calculated, the total time of the longest video segment is 6.52s, so the threshold T_{max} of the capture time can be set to 7s.

F. EXTRACTION OF BATTING ACTION

1) SLIDING WINDOW EXTRACTION

Sliding window segmentation plays a key role in extracting badminton hitting movements. By dividing the continuous video frame sequence into small windows, this method is helpful to focus on the local area and capture the key features of the hitting action. This includes the interaction between the racket and badminton, the movements of arms and bodies, and the trajectory of the ball. Meanwhile, the sliding window method also creates time series data, which enables people to analyze the time evolution of hitting action, including the beginning, middle and end stages. In this way, the transition between different hitting actions is easier to detect and segment, which is helpful to accurately extract and analyze each action. Each sliding window can be used to extract various visual features, such as position, speed and acceleration, which are helpful to classify and identify different types of hitting actions. In a word, the sliding window segmentation method improves the accuracy and efficiency of action recognition, which is very important for badminton technical analysis and athlete training.

The big data division technology based on the sliding window is to divide the collected shot information into several windows of the same width. The adjacent windows can overlap with each other or even do not intersect. This section will take the action of players pushing and hitting the ball with the forehand many times in continuous confrontation as an example to explore the division technology of sliding window.

Sliding windows usually use a small window of fixed length to move along the time axis and achieve the acquisition of hitting dynamics. The frequency band of data collection is usually 200Hz, and the badminton hitting dynamics is usually not greater than 0.5s. Therefore, the small window width is set to 100 to gain the hitting dynamics. Equations (9)-(10) are used to estimate its net resultant acceleration and complete the batting action extraction.

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2 - 1} \tag{9}$$

In (9), a_x, a_y, a_z are velocity signals representing three axes, and -1 is the gravity velocity fraction that is always retained. a is the net resultant acceleration that characterizes pure human action, and the unit of acceleration rate is g.

$$A_m = \underbrace{arg\,max \left[\frac{1}{k} \sum\nolimits_m^{m+k-1} a_m^2 \right]}_{m} \tag{10}$$

In (10), a_m represents the net resultant acceleration at the mth sample point. k is the window width, and 100 is taken here. A_m represents the maximum averaged sum of squares of net resultant accelerations with a window width of 100.

Sliding window segmentation technology has a series of advantages in analyzing badminton hitting action. Firstly, it allows the effective extraction of dynamic features. By dividing the continuous hitting information into windows, people can better capture and understand the dynamic changes of hitting actions. Secondly, the width of the sliding window is adjustable, which makes it suitable for different types of hitting actions and time scales. In addition, the technology also provides a method of feature extraction, such as net acceleration and other related features, which can be used for further analysis and identification and help to further study the nature of hitting action.

However, the sliding window segmentation technology also has some limitations. Firstly, the selection of window width is very important, and improper selection may lead to the loss of information or the introduction of noise. Secondly, the data acquisition frequency is very important for the accuracy of the analysis, and the lower acquisition frequency may not be suitable for some high-speed hitting actions. Finally, the nature of different hitting actions is different, and the applicability of sliding window segmentation technology varies with the type of action, which may require different analysis methods to deal with different situations. Therefore, when applying this technology, it is necessary to carefully select the window width and consider the characteristics of data acquisition frequency and action type.

2) EVENT WINDOW EXTRACTION

Event window segmentation is a commonly used technique in motion analysis, which is used to identify and extract specific events or actions. In badminton action recognition, it is used to capture and analyze the hitting action. The core idea of this method is to divide the sports data into different time windows according to the occurrence time of specific events to analyze and identify these events. In badminton, the event window can be used to capture the start, end and different stages of hitting the ball, and the start and end times of the event window can be determined by identifying the characteristics or signs of specific events. This technique is helpful to extract the key information of hitting action, such as timing, speed and strength, and is very helpful for the technical analysis and improvement of athletes and real-time training feedback.

In a word, event window segmentation is of great significance to the technical improvement and training of badminton.

The event-based window data division technique divides a continuous shot data into several windows with different widths. The start time and end time of each window represent the start and end time of the event, and the start and end time of the hitting action corresponding to the badminton refers to the player's starting time and hitting time. In the process of hitting the ball, the acceleration will increase greatly when the player starts the shot, so the start time of the shot can be judged by fitting the inverse process of the acceleration signal. Besides, whether the acceleration rate of increase exceeds a certain rate can also be determined. If the acceleration at impact has reached its peak throughout the swing, the peak can be measured in the neighborhood to confirm the impact.

In badminton action recognition, using event window segmentation technology faces a series of challenges. Firstly, it is very important for accurate identification to accurately determine the occurrence time of hitting events. However, due to the interference of noise and different marking methods, the accuracy of event location is a challenge. Secondly, the diversity and complexity of badminton movements may lead to similar characteristics of different events, which increases the difficulty of identification. In addition, the real-time requirement means that the algorithm needs to complete motion recognition in a short time, so speed and efficiency are also test points. In addition, data quality, data volume and labeling cost are also challenges, and a large number of high-quality data and labeling are necessary conditions for training and verifying the model. The identification of athletes in multi-person and multi-scene situations and the robustness of the algorithm also need to be considered. To sum up, badminton action recognition using event window segmentation needs to deal with many challenges such as positioning, diversity, real-time, data quality, data volume, multi-person and multi-scene, algorithm robustness and so on.

G. ACTION FRAMING AND FEATURE SELECTION

Clustering eigenvalue plays a key role in badminton hitting action analysis. They are used to classify, identify and analyze different types of badminton strokes. Firstly, the clustering eigenvalue can group similar hitting actions, which helps to clearly distinguish different types of actions such as forehand picking, forehand pushing and backhand hitting. Secondly, by comparing the observed action characteristics with the cluster eigenvalues of known categories, the system can accurately identify the athletes' hitting actions in real-time or offline mode. In addition, the cluster eigenvalue also provides a quantitative analysis method of hitting action, which can be used to evaluate the quality, speed, strength and other parameters of the action, and help coaches and athletes understand the technical performance and provide suggestions for improvement. The most important thing is that the calculation of cluster eigenvalues is usually real-time, which enables the system to respond quickly when athletes hit the ball and provide timely feedback and suggestions, thus improving the athletes' technical level and training effect. Therefore, cluster eigenvalue plays an indispensable role in badminton hitting action analysis.

Assuming that all ball dynamics are decomposed into n element dynamics, the statistic of each element dynamics is 100/n. According to the empirical value, n=10 is taken in this paper. All ball dynamics are solved into ten meta dynamics. It is also a fixed-length waveform for each meta action. The feature evaluation by cluster analysis method can be performed on the waveform feature value, or two feature values with large cluster feature values can be selected as evaluation parameters. Here, the clustering feature amounts of the same feature amount can be evaluated by (11)-(13).

$$\lambda_{i,j} = \frac{x_{i,j}}{\sum_{j=1}^{k} x_{i,j}} \tag{11}$$

$$\lambda_j = \frac{\sum_{i=1}^{k} \lambda_{i,j} x_{i,j}}{\sum_{i=1}^{k} x_{i,j}} \tag{12}$$

$$\eta = \frac{1}{k} \sum_{j=1}^{k} \lambda_j \tag{13}$$

k is the total number of batting action classes. $x_{i,j}$ represents the total number of the i-th dynamic identified as the j-th dynamic, so the denominator in (11) represents the total number of the i-th dynamic. $\lambda_{i,j}$, the transfer coefficient, represents the weight of the i-th type of action being recognized as the j-th type. The numerator in (12) represents the number of samples identified as the jth class after being weighted by the transfer coefficient. λ_j represents the probability that the sample set is clustered into action j, and it is also called the clustering characteristic value of action j. In (13), η represents the mean cluster analysis eigenvalue. The corresponding sample mean value and cluster analysis method characteristics under the average feature value in different time domains and frequency intervals are calculated by taking the average feature value of each frame segment of the divided sample data.

The average value and peak value play an important role in the feature selection of badminton hitting action. The average value is used to indicate the average strength, speed or position of the action, while the peak-to-peak value reflects the maximum strength or speed in the action. These two statistical characteristics provide important information about the quality and stability of the action. For example, the average value can be used to describe the average speed of picking the ball, while the peak value can represent the maximum speed or strength of backhand hitting. By comparing the average value and peak value of different actions, the differences between them can be identified for the classification and recognition of actions.

In addition, these characteristics can also evaluate the technical level of athletes, monitor the stability of movements, and provide guidance for coaches to improve and optimize movements.

H. VECTOR QUANTIZATION AND IMPROVEMENT OF BAUM-WELCH TRAINING ALGORITHM

1) VECTOR QUANTIZATION

The purpose of vector quantization is to represent high-dimensional data as lower-dimensional vectors for data analysis, comparison and processing. The purpose of vector quantization in this paper is to express the meta-dynamic (sub-action) of badminton hitting action as a 6-dimensional feature vector containing peak-to-peak value and average value information. Such a digital representation is helpful to transform the hitting action into a form that can be processed by computer, which makes the subsequent analysis and recognition more efficient and accurate. Vector quantization has an important influence on the analysis of hitting action, because it extracts the key features of the action and expresses them in digital form. In this way, different hitting actions can be compared, clustered and analyzed to better understand and identify different hitting actions. This digital representation can also be used to train and optimize the model, which improves the accuracy and robustness of hitting action recognition. Therefore, vector quantization plays a key role in badminton action analysis.

The two largest characteristic quantities in the clustering eigenvalues - the peak-to-peak value and the average number are selected as the characteristic quantities of the meta-dynamics. Therefore, all meta-dynamics can be represented as a six-dimensional feature vector including the peak-to-peak value and average value of the X, Y, and Z axes. An unsupervised learning clustering algorithm can be used to perform vector digitization on the characteristic vector of each meta-dynamic. After a secondary clustering operation is completed, vector digitization is implemented on each element dynamics of the sample, as shown in (14)-(15).

$$\mu_i = \frac{1}{N} \sum\nolimits_{x_j \in s_i} x_j \tag{14}$$

$$J = \underbrace{argmin}_{S} \sum\nolimits_{i=1}^{k} \sum\nolimits_{x_j \in S_i} \left\| x_j - \mu_i \right\|^2 \tag{15}$$

x_j is the sample data. $S = \{S_1, S_2, \ldots, S_k\}$ represents the k spaces divided by the sample. μ_i is the average value of the sample in the category S_i, which is the cluster center point. J represents the cluster quantization error function.

As a result, the corresponding relationship between the value of k and the quantization error is obtained by dynamically adjusting the value of k and evaluating its corresponding clustering error and algorithm time complexity.

In vector quantization, different values of k (the number of clustering centers) will affect the number of iterations and clustering errors. A larger value of k usually requires more iterations to converge to the best clustering center, and increasing the value of k may lead to an increase in clustering error, because the data becomes more complicated. Therefore, when choosing the value of k, people need to make a trade-off to avoid under-fitting and over-fitting. Cross-validation or elbow rule can be used to select the appropriate k value to find a balance between the number of iterations and the clustering error.

2) IMPROVEMENT OF THE BAUM-WELCH TRAINING ALGORITHM

The data analysis of each hitting action in the sample can be collected into several samples, and the sample used in the training HMM parameter modeling is a single sample, so the data analysis in the modeling may enter the local optimal estimation. The recognition rate is low for some samples. Regarding multi-sample training, two methods are given. The first is the average practice method. According to the input of the model, the training statistics of multiple groups in the same model are averaged to complete the model practice after the data layer fusion is performed on the observation sequence of the samples. The second is the frequency weight practice method. The Baum-Welch training algorithm is improved after applying linear weighting to the sample observation order and the frequency of events that occur in this mode.

3) EXPERIMENTAL DESIGN OF HMM ACTION RECOGNITION BASED ON DIFFERENT PREPROCESSING ALGORITHMS

Experimental purpose: the three different separation methods are compared through the data separation technology provided by the batting action in the data preprocessing calculation to select the preprocessing algorithm with the high recognition rate.

Option 1: segmentation is performed by a sliding window when hitting the ball. For each hitting point, a peak test of hitting time is made, and fifty sample points around the hitting time will become the dynamic characteristics of hitting.

Option 2: segmentation is performed using a sliding window with a maximum value. Selecting a sliding window with a window length of one hundred allows for maximum checking and resetting. The generation of each maximum value is relative to the progress of a batting action.

Option 3: event-based window segmentation. The two times between the player's starting and hitting will be detected and used as the window segmentation of the starting and ending time points.

Experimental method: the univariate principle is implemented. The three different hitting motion extraction algorithms described above are used, while other methods of statistics, practice, and identification remain unchanged. The HMM exercises are completed using similar sample data results, and the identification of the six hitting sports is completed through the same experimental data results.

Different preprocessing methods have influenced the recognition rate of HMM in badminton action recognition, which can provide valuable enlightenment for intelligent data analysis and auxiliary training products of badminton. By using different preprocessing methods, such as sliding window segmentation, maximum sliding window segmentation and event-based window segmentation, the recognition performance of HMM's hitting action can be significantly affected. Different preprocessing methods may lead to different feature extraction and data separation effects, and then affect the accuracy of HMM model. By comparing these methods through experiments, people can determine which method is better in improving the recognition rate of badminton movements.

By comparing the performance of different preprocessing algorithms, it can provide guidance for intelligent data analysis and auxiliary training products of badminton. Specifically, the preprocessing algorithm with high recognition rate can become a part of the intelligent data analysis platform to help professional players and coaches better analyze and improve their hitting movements. This can improve the training effect, help players give full play to their personal advantages and overcome their shortcomings, and promote the intelligent development of badminton. Therefore, by comparing the performance of different preprocessing algorithms, people can determine the most suitable method for badminton action recognition, improve the recognition rate, provide better training support for athletes and coaches, and promote the improvement of badminton technology and training level.

Fig. 5 depicts the experimental design process for HMM action recognition based on different preprocessing algorithms.

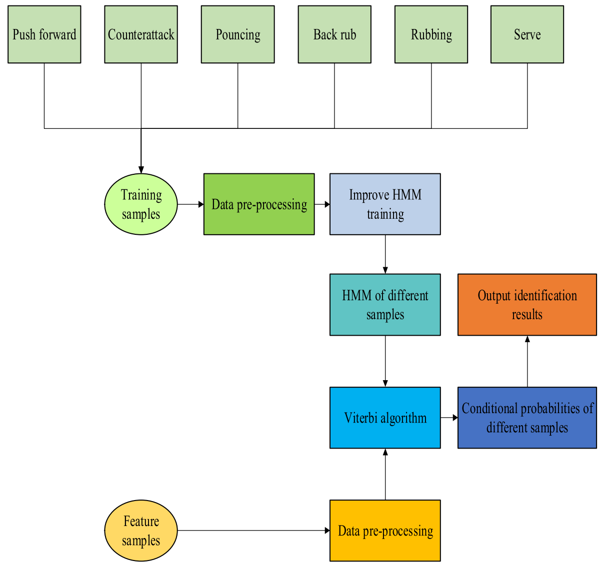


FIGURE 5. The experimental design process of HMM action recognition based on different preprocessing algorithms.

III. RESULTS OF MODEL TEST

A. COMPARISON BETWEEN THE ORIGINAL RESNET MODE AND THE RP-RESNET MODE

The following experimental comparisons are carried out to enrich the experimental data and compare and verify the possibility of the RP-ResNet network in depth. The detection efficiency comparison between the original ResNet mode and the RP-ResNet mode is revealed in Fig. 6.

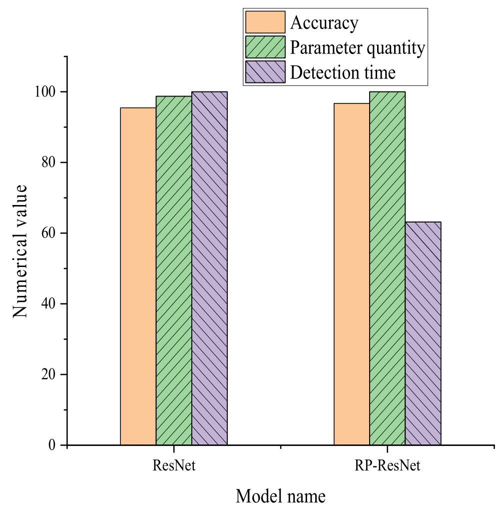
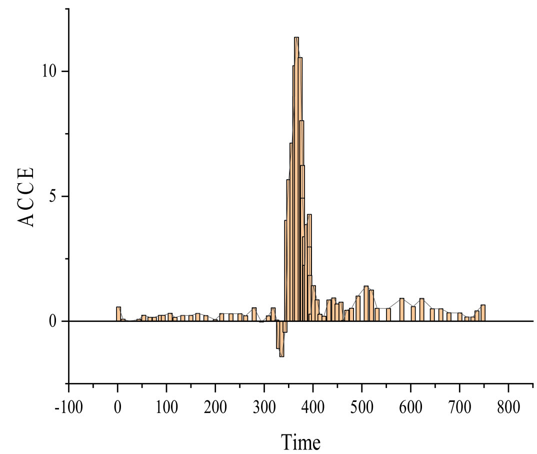


FIGURE 6. Comparison of experimental results between the original ResNet model and RP-ResNet.

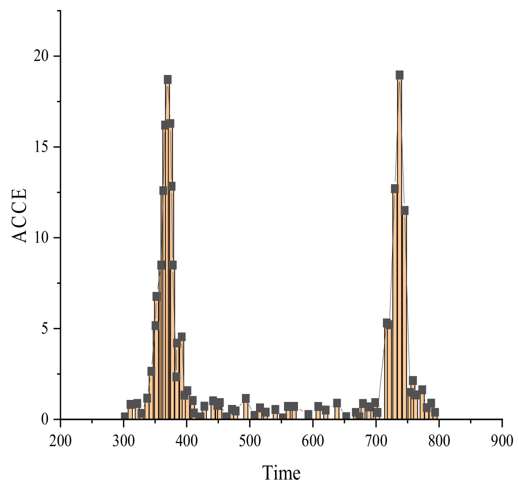
From Fig. 6, it can be found that the RP-ResNet network model has obvious advantages in all aspects, compared with the original model ResNet. The accuracy of RP-ResNet increased by 1.25%, reaching 96.72%. This means that RP-ResNet can identify and classify the dynamic information in badminton matches more effectively, which is very beneficial for tasks with high accuracy requirements. Although the number of parameters of RP-ResNet increases by 0.1% compared with the original model, this slight increase does not sacrifice the performance of the model. This shows that RP-ResNet makes more efficient use of parameters and achieves better cost performance. The detection time of RP-ResNet is significantly reduced by 0.07 unit. This means that RP-ResNet can realize the real-time identification and analysis of badminton game dynamic information more quickly, which provides higher real-time performance and efficiency. Generally speaking, RP-ResNet network model has achieved remarkable success in improving accuracy, optimizing the number of parameters and reducing detection time. These improvements make RP-ResNet a higher performance and higher efficiency choice, and provide badminton enthusiasts with better technical statistical analysis tools, which is expected to promote the automation and scientific development of badminton competition.

B. HITTING ACTION EXTRACTION BASED ON SLIDING WINDOW SEGMENTATION

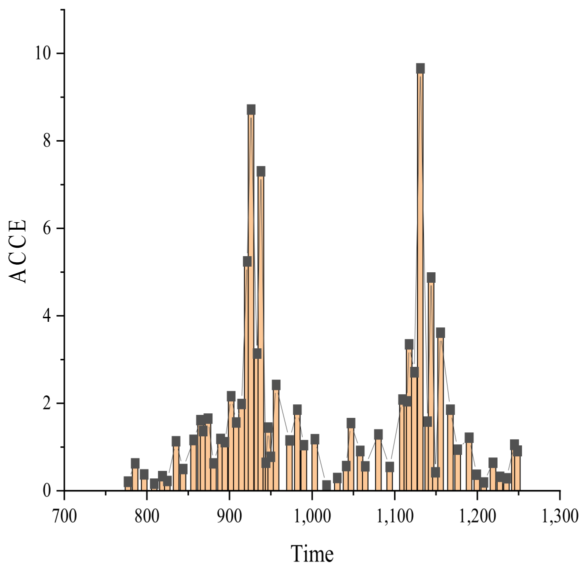
The minimum threshold set is 3, and a maximum reset can be done to facilitate multiple shots in a row. Fig. 7(a) shows the schematic diagram of the net resultant acceleration. Fig. 7(b)



(a) Calculation result of the net resultant acceleration



(b) Schematic diagram of sliding window segmentation for forehand picking



(c) Schematic diagram of forehand push sliding window segmentation

FIGURE 7. Shot action extraction for sliding window segmentation.

and Fig. 7(c) demonstrate schematic diagrams of sliding window segmentation for forehand pick and forehand push based on net resultant acceleration information. The staggered axis is the sampling time axis, and the vertical axis Accelerometer is the net resultant acceleration information. The peak-to-peak part is the statistical information of the forehand pick and forehand push in the sliding window.

In Fig. 7, firstly, the calculated results of net combined acceleration show the acceleration values at different time points, reflecting the acceleration and deceleration of athletes in different hitting actions. Secondly, the calculation results of forehand ball-picking sliding show the sliding situation of athletes in forehand ball-picking, and the sliding results at different time points may be related to the strength and skills of ball-picking. Finally, the sliding calculation results of forehand push show the sliding situation of athletes in forehand push. Similarly, the sliding calculation results at different time points may be related to the strength and control of push. These data provide detailed action information, which can help to understand the characteristics and trends of different actions and their relationship with technical level and competition strategy through in-depth analysis, which is of great value to athletes' training and performance evaluation.

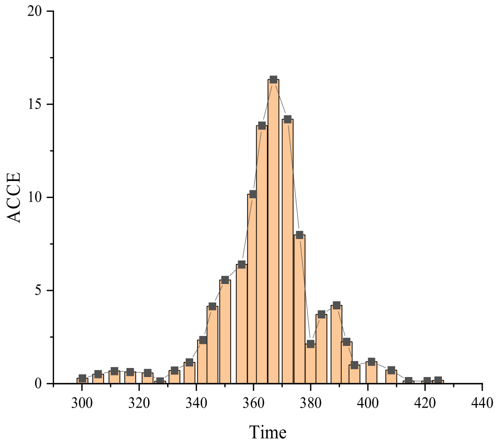
C. HITTING ACTION EXTRACTION BASED ON EVENT WINDOW SEGMENTATION

Fig. 8(a) and Fig. 8(b) are schematic diagrams of the event window segmentation of forehand pick and forehand push using net resultant acceleration information, respectively. The staggered axis is the sampling time axis, and the vertical axis is the net resultant acceleration information. The peak value represents the measured starting point and hitting location, and the peak-to-peak value part is the information about the forehand pick and forehand putt hitting motion provided by the event window.

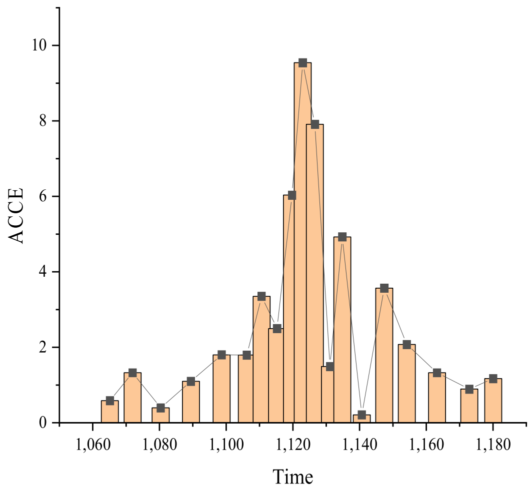
In Fig. 8, the event values of forehand picking and forehand pushing events change obviously at different time points, which may reflect the differences in strength and skills of athletes in different hitting actions. However, there are some missing event values in the data, which may affect the subsequent analysis and identification. In addition, it is pointed out that the results of event-based window extraction do not completely include the acceleration information before hitting the ball. The obtained numerical widths are inconsistent, so time warping is needed to solve these problems to ensure the recognition rate and analysis accuracy of the system. Generally speaking, these data provide important information for a deeper understanding of athletes' movements and technical level, but further processing is needed to ensure the integrity and availability of the data.

D. ANALYSIS OF ACTION FRAMING AND FEATURE SELECTION RESULTS

Fig. 9 shows the results of clustering characteristic values of common batting actions.



(a) Schematic diagram of window segmentation for forehand picking event



(b) Schematic diagram of forehand push event window segmentation

FIGURE 8. **Shot action extraction based on event window segmentation.**

In Fig. 9, the peak-to-peak value is 0.175, which shows that there is a great difference in amplitude between different actions, and there may be a great difference in actions. The variance is 0.152, which shows that the characteristics of these actions are relatively stable and consistent statistically. The coefficient of Fast Fourier Transform (FFT) is 0.15, which indicates that the features in frequency domain are also clustered. The energy is 0.143, which may indicate that the energy distribution of actions has certain similarity. The spectral density is 0.142, which indicates that the features in frequency domain are also clustered. The entropy is 0.125, which may indicate that the diversity of information in the sample data is relatively small. Generally speaking, the statistical results of these data provide important clues for understanding the characteristics and similarities of different hitting actions, especially the average value and peak-peak value, which show good clustering characteristics and may

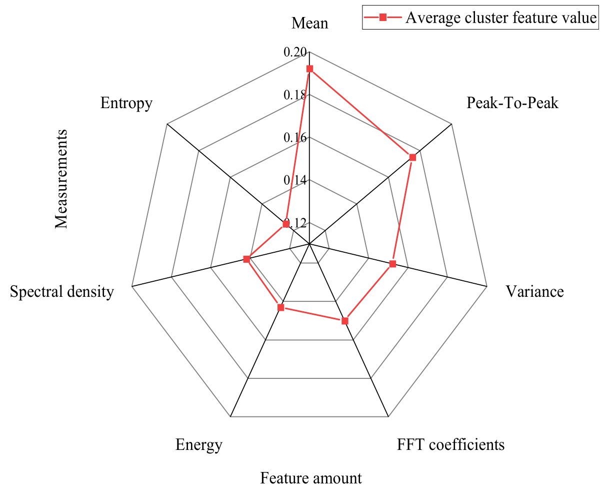


FIGURE 9. **Comparison of the average clustering characteristic values of various feature quantities.**

be valuable in action classification and identification. These findings are helpful to deeply understand the differences and commonness among hitting actions, and are of guiding significance for further action analysis and technical improvement.

E. VECTOR QUANTIZATION EVALUATION

The change curve of the number of iterations and the clustering error with the value of k is demonstrated in Fig. 10.

In Fig. 10, when k=7, the clustering error is 58.88556, and the number of iterations is 34.31489. When k=15, the clustering error is 47.84302 and the number of iterations is 26.27138. When k=24, the clustering error is 61.37162 and the number of iterations is 20.3707. When k=32, the clustering error is 33.04263 and the number of iterations is 25.04273. When k=40, the clustering error is 38.71465 and the number of iterations is 18.42918. When k=48, the clustering error is 45.81335 and the number of iterations is 13.24324. When k=55, the clustering error is 26.05616 and the number of iterations is 16.34386. When k=64, the clustering error is 30.5858 and the number of iterations is 22.01494. When k=72, the clustering error is 35.54307 and the number of iterations is 10.40004. When k=80, the clustering error is 20.78535 and the number of iterations is 16.78587. By comprehensive analysis, it shows that with the increase of K value, the clustering error shows a downward trend, and the change of iteration times is relatively complicated. When k=32, the clustering error is low and the number of iterations is relatively small, which shows that the clustering effect is good and the calculation efficiency is high at this point. These data are helpful for people to choose the appropriate K value for clustering analysis to achieve the best clustering effect.

F. HMM ACTION RECOGNITION EXPERIMENTS BASED ON DIFFERENT PREPROCESSING ALGORITHMS

The action recognition rates of HMM for different preprocessing methods are revealed in Fig. 11.

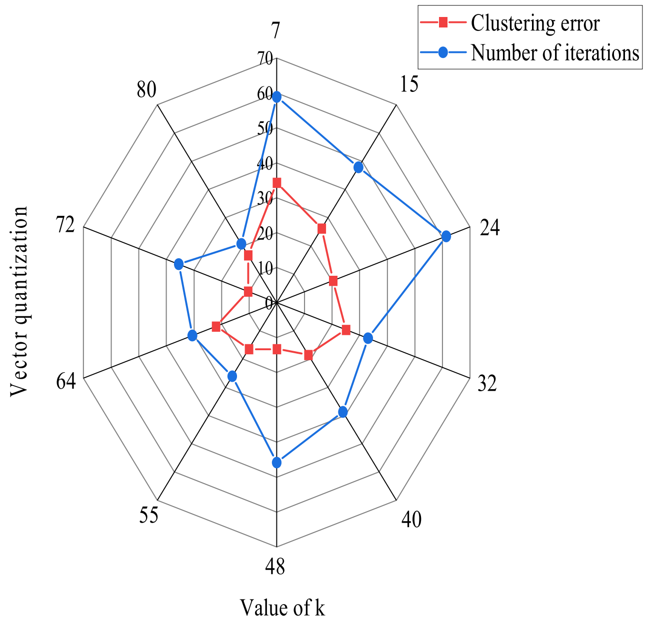


FIGURE 10. k-means vector quantization characteristic evaluation diagram.

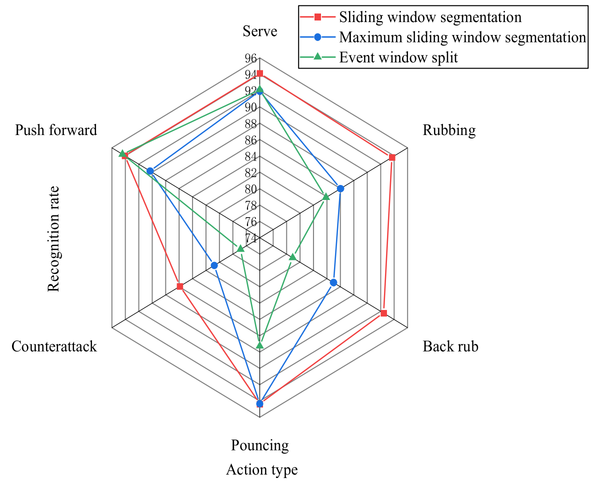


FIGURE 11. Action recognition rate of HMM with different preprocessing methods.

In Fig. 11, three completely different bar-shaped color blocks represent the above three completely different acquisition methods of hitting dynamics, the horizontal axis represents the above six completely different hitting dynamics, and the vertical axis represents the discrimination rate of various hitting dynamics. In different hitting actions, the scores of different preprocessing methods are different: the score of Serve's sliding window segmentation is 94.09338, the maximum sliding window segmentation is 91.92832 and the event window segmentation is 92.11599. Rubbing's sliding window segmentation score is 93.65307, the maximum sliding window segmentation is 86.01075, and the event window segmentation is 83.84642. Back rub's sliding window segmentation score is 92.42705, the maximum sliding window segmentation is 84.98116 and the event window segmentation is 78.90294. Pouncing's sliding window segmentation score is 94.33873, the maximum sliding window segmentation is 94.32631 and the event window segmentation is 87.27035. Counterattack's sliding window segmentation score is 85.8742, the maximum sliding window segmentation is 80.77591, and the event window segmentation is 76.85107. The sliding window segmentation score of Push forward is 94.04445, the maximum sliding window segmentation is 90.31384, and the event window segmentation is 94.41248. Generally speaking, the first scheme has a slightly higher determination rate of hitting action.

IV. DISCUSSION

This paper significantly differs from previous research methods. Previous research primarily relied on traditional methods of manual instruction, including one-on-one or one-to-many training. While this approach emphasizes individual skill training, it has several limitations. First, it often fails to provide highly personalized training, thus unable to fully leverage an athlete's individual strengths and address their weaknesses. Next, previous research methods are constrained in terms of technical analysis and data collection, lacking the capability for automation and real-time feedback. Most importantly, this approach is ill-suited to meet the growing demands in the field of sports and fitness, especially with the increasing popularity of badminton.

In contrast, this paper introduces innovative technologies such as deep learning and binocular vision systems. These technologies offer the advantages of advanced action recognition, enabling automated training and real-time feedback. The binocular vision system accurately captures the three-dimensional position of the shuttlecock on the court, facilitating advanced motion analysis. Additionally, using deep learning models such as the convolutional neural network can efficiently identify the players' hitting actions, enhancing measurement accuracy, especially for small target objects. The efficiency of network features is further improved through an improved pyramid pooling structure and SENet mechanism, which increases measurement accuracy for small objects. The research method can advance badminton sports technology and intelligence, providing athletes with better support and training opportunities. It addresses the shortcomings of previous research methods and offers more technical support for sports enthusiasts and professional athletes. This contributes to the automation of badminton training and technological advancement. The exact research significance lies in addressing the challenge of obtaining professional guidance, which has become increasingly relevant due to the growing focus on health, especially the popularity of badminton as a fitness activity. Furthermore, with technological advancements, this paper introduces an innovative artificial intelligence platform for badminton sports, tackling various technical challenges in both hardware and software engineering, ranging from data collection to computational programming.

V. CONCLUSION

This paper presents an algorithm for recognizing dynamic information in badminton matches through a single sensor. It establishes an application platform for real-time technical statistical analysis of badminton matches. For badminton enthusiasts, a feasible method is proposed to accelerate their skill development. Experimental results demonstrate that the deep neural network's lower-level network can generate clear target positioning signals for initial target localization. The embedded SENet mechanism widens the observation angle of feature maps and strengthens connections between channels. The improved spatial pyramid structure enhances the recognition of details for both small and large targets. The experiments show that a sliding window can achieve a high dynamic recognition rate, and extracting features for all strokes through a sliding window can yield favorable results. Peak checks are performed on the hit time for each stroke, and fifty sample points around the hit time are taken as dynamic hit features.

This paper introduces an algorithm for recognizing dynamic information in badminton matches using a single sensor. It aims to help badminton enthusiasts improve their skill levels. This method is user-friendly, achieves accurate target localization through deep neural networks, enhances target detail recognition, and boasts a high dynamic recognition rate. This innovative tool offers technical support for badminton players and enthusiasts, with the potential to advance the automation and popularity of badminton.

REFERENCES

[1] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," *WIREs Data Mining Knowl. Discovery*, vol. 9, no. 4, p. 1312, Jul. 2019.

[2] M. ul Hassan and A. A. Aziz, "Investigating the use of computer technology for E-learning in Pakistani maddaris: Case of religious teachers," *Int. J. Distance Educ. E-Learn.*, vol. 5, no. 1, pp. 44–56, 2019.

[3] S. A. Umpleby, T. A. Medvedeva, and V. Lepskiy, "Recent developments in cybernetics, from cognition to social systems," *Cybern. Syst.*, vol. 50, no. 4, pp. 367–382, May 2019.

[4] J. Gauthier, A. T. Vincent, S. J. Charette, and N. Derome, "A brief history of bioinformatics," *Briefings Bioinf.*, vol. 20, no. 6, pp. 1981–1996, Nov. 2019.

[5] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, and L. Zdeborová, "Machine learning and the physical sciences," *Rev. Modern Phys.*, vol. 91, no. 4, 2019, Art. no. 045002.

[6] S. Garg, N. Sünderhauf, F. Dayoub, D. Morrison, A. Cosgun, G. Carneiro, Q. Wu, T.-J. Chin, I. Reid, S. Gould, P. Corke, and M. Milford, "Semantics for robotic mapping, perception and interaction: A survey," *Found. Trends Robot.*, vol. 8, nos. 1–2, pp. 1–224, 2020.

[7] S. Schneider, G. W. Taylor, S. Linquist, and S. C. Kremer, "Past, present and future approaches using computer vision for animal re-identification from camera trap data," *Methods Ecol. Evol.*, vol, vol., 10, no. 4, pp. 461–470, 2019.

[8] K. Wang, "A smart robot training data acquisition and learning process recording system based on blockchain," *Open Access Library J.*, vol. 7, no. 9, pp. 1–5, 2020.

[9] Y. Zhao, J. Li, and J. E. Wang, "Analysis and prediction of 'AI+ education' attention based on Baidu index—Taking Guizhou province as an example," *Future Internet*, vol. 13, no. 5, p. 120, 2021.

[10] H. Zheng and W. He, "Relevant concepts, key technologies and typical application scenarios of AI + education," *J. Phys., Conf. Ser.*, vol. 1861, no. 1, Mar. 2021, Art. no. 012059.

[11] S. Mori, K. Tanaka, S. Nishikawa, R. Niiyama, and Y. Kuniyoshi, "High-speed humanoid robot arm for badminton using pneumatic-electric hybrid actuators," *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 3601–3608, Oct. 2019.

[12] M.-H. Hung, C.-Y. Chang, K.-C. Lin, C.-L. Hung, and C.-S. Ho, "The applications of landing strategies in badminton footwork training on a backhand side lateral jump smash," *J. Hum. Kinetics*, vol. 73, no. 1, pp. 19–31, Jul. 2020.

[13] A. Mathis, S. Schneider, J. Lauer, and M. W. Mathis, "A primer on motion capture with deep learning: Principles, pitfalls, and perspectives," *Neuron*, vol. 108, no. 1, pp. 44–65, Oct. 2020.

[14] H.-B. Zhang, Y.-X. Zhang, B. Zhong, Q. Lei, L. Yang, J.-X. Du, and D.-S. Chen, "A comprehensive survey of vision-based human action recognition methods," *Sensors*, vol. 19, no. 5, p. 1005, Feb. 2019.

[15] N. S. Logan, H. Radhakrishnan, F. E. Cruickshank, P. M. Allen, P. K. Bandela, L. N. Davies, S. Hasebe, S. Khanal, K. L. Schmid, F. A. Vera-Diaz, and J. S. Wolffsohn, "IMI accommodation and binocular vision in myopia development and progression," *Investigative Opthalmol. Vis. Sci.*, vol. 62, no. 5, p. 4, Apr. 2021.

[16] F. Mueller, M. Davis, F. Bernard, O. Sotnychenko, M. Verschoor, M. A. Otaduy, D. Casas, and C. Theobalt, "Real-time pose and shape reconstruction of two interacting hands with a single depth camera," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 1–13, Aug. 2019.

[17] H. Noble and R. Heale, "Triangulation in research, with examples," *Evidence Based Nursing*, vol. 22, no. 3, pp. 67–68, Jul. 2019.

[18] R. Giordano-Santini, E. Kaulich, K. M. Galbraith, F. K. Ritchie, W. Wang, Z. Li, and M. A. Hilliard, "Fusogen-mediated neuron-neuron fusion disrupts neural circuit connectivity and alters animal behavior," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 37, pp. 23054–23065, Sep. 2020.

[19] J.-C. Yin, L. Zhang, N.-X. Ma, Y. Wang, G. Lee, X.-Y. Hou, Z.-F. Lei, F.-Y. Zhang, F.-P. Dong, G.-Y. Wu, and G. Chen, "Chemical conversion of human fetal astrocytes into neurons through modulation of multiple signaling pathways," *Stem Cell Rep.*, vol. 12, no. 3, pp. 488–501, Mar. 2019.

[20] D. Zou, Y. Cao, D. Zhou, and Q. Gu, "Gradient descent optimizes over-parameterized deep ReLU networks," *Mach. Learn.*, vol. 109, no. 3, pp. 467–492, Mar. 2020.

[21] G. W. Lindsay, "Convolutional neural networks as a model of the visual system: Past, present, and future," *J. Cognit. Neurosci.*, vol. 33, no. 10, pp. 2017–2031, Sep. 2021.

[22] Z. Li, S. Wang, R. Fan, G. Cao, Y. Zhang, and T. Guo, "Teeth category classification via seven-layer deep convolutional neural network with max pooling and global average pooling," *Int. J. Imag. Syst. Technol.*, vol. 29, no. 4, pp. 577–583, Dec. 2019.

[23] Y. Zhang, Y. Hou, S. Zhou, and K. Ouyang, "Encoding time series as multi-scale signed recurrence plots for classification using fully convolutional networks," *Sensors*, vol. 20, no. 14, p. 3818, Jul. 2020.

[24] M. Gao and G.-Y. Shi, "Ship spatiotemporal key feature point online extraction based on AIS multi-sensor data using an improved sliding window algorithm," *Sensors*, vol. 19, no. 12, p. 2706, Jun. 2019.

[25] W.-S. Hu, H.-C. Li, L. Pan, W. Li, R. Tao, and Q. Du, "Spatial–spectral feature extraction via deep ConvLSTM neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4237–4250, Jun. 2020.

[26] F. Li and H. Yu, "An investigation of single-domain and multidomain medication and adverse drug event relation extraction from electronic health record notes using advanced deep learning models," *J. Amer. Med. Inform. Assoc.*, vol. 26, no. 7, pp. 646–654, Jul. 2019.

[27] J. Reinecke and S. Ansari, "Microfoundations of framing: The interactional production of collective action frames in the occupy movement," *Acad. Manage. J.*, vol. 64, no. 2, pp. 378–408, Apr. 2021.

[28] V. Rossa-Roccor, A. Giang, and P. Kershaw, "Framing climate change as a human health issue: Enough to tip the scale in climate policy?" *Lancet Planet. Health*, vol. 5, no. 8, pp. e553–e559, Aug. 2021.

[29] I. M. Khater, I. R. Nabi, and G. Hamarneh, "A review of super-resolution single-molecule localization microscopy cluster analysis and quantification methods," *Patterns*, vol. 1, no. 3, Jun. 2020, Art. no. 100038.

[30] S. B. U. D. Tahir, A. Jalal, and K. Kim, "Wearable inertial sensors for daily activity analysis based on Adam optimization and the maximum entropy Markov model," *Entropy*, vol. 22, no. 5, p. 579, May 2020.

[31] M. Sesia, C. Sabatti, and E. J. Candès, "Gene hunting with hidden Markov model knockoffs," *Biometrika*, vol. 106, no. 1, pp. 1–18, Mar. 2019.