

Research Article

Research on the Identification and Evaluation of Aerobics Movements Based on Deep Learning

Hua Zhao,¹ Aibo Wang,² and Ying Fan ³

¹Department of Sport Arts, Hebei Sport University, Shijiazhuang 050041, Hebei, China

²Department of Physical Training, Hengshui University, Hengshui 053000, Hebei, China

³School of Arts, Beijing Sport University, Beijing 100084, China

Correspondence should be addressed to Ying Fan; fy1121@st.btbu.edu.cn

Received 21 October 2021; Revised 23 November 2021; Accepted 24 November 2021; Published 21 December 2021

Academic Editor: Tongguang Ni

Copyright © 2021 Hua Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A deep learning approach is used in this study to provide insight into aerobics movement recognition, and the model is used for aerobics movement recognition. The model complexity is significantly reduced, while the multi-scale features of the target at the fine-grained level are extracted, significantly improving the characterization of the target, by embedding lightweight multi-scale convolution modules in 3D convolutional residual networks to increase the local perceptual field range in each layer of the network. Finally, using the channel attention mechanism, the key features are extracted from the multi-scale features. To create a dual-speed frame rate detection model, the fast-slow combination idea is fused into a 3D convolutional network. To obtain spatial semantic information and motion information in the video, the model uses different frame rates, and the two-channel information is fused with features using lateral concatenation. Following the acquisition of all features, the features are fed into a temporal detection network to identify temporal actions and to design a behavior recognition system for the network model to demonstrate the network model's applicability. The average scores of students in the experimental group were significantly higher than those in the control group in seven areas: set accuracy, movement amplitude, movement strength, body coordination, coordination of movement and music, movement expression, and aesthetics; the average scores of movement proficiency and body control in the experimental group were also significantly higher than those in the control group, but the differences were not significant. The differences between the eight indicators in the experimental group were not significant when compared to those in the preexperimental group, indicating that intensive rhythm training for students improves secondary school students' comprehension, proficiency, and presentation of aerobics sets.

1. Introduction

After the development of aerobics as an international competition, the competition structure has been continuously differentiated, and trying to highlight the competitive characteristics, the rules as the project development of the wind vane, is the best embodiment of the project competitive characteristics. There is a new interpretation of the assessment of the athletes' participation in the sets, and the excellent results not only rely on high technical ability but also require the all-around development of athletic ability, especially the need to have excellent physical fitness [1]. Although aerobics is a skill-driven event group, strong

physical fitness is essential to complete the high standard of the competition within the stipulated competition time. Good physical fitness not only can drive the improvement of skills, but also is a necessary condition to complete the competition goals and improve the competition results; excellent strength quality is to achieve the guarantee of solid improvement of difficult techniques, superendurance quality helps the competition in high-quality play, and the combination of movement and fast and powerful movements helps to improve the artistic score. In aerobics, as in diving, competitive gymnastics, wushu, and other sports, the judges assess the athletes' performance subjectively according to the scoring rules [2]. The control group adopts

the traditional aerobics teaching methods. Teachers pay more attention to the training of various technical movement specifications and routines of aerobics when teaching. Music is an indispensable part of aerobics. At present, for athletes in such sports, the identification of their athletic ability and athletic rating is often based on the results achieved by the athletes in various competitions. This is inevitably influenced by the size of the competition and the level of the participants. Therefore, it is very meaningful to establish a set assessment system for athletes' specific athletic competitive abilities in these sports. With this assessment system, coaches can use a unified quantitative standard to assess the level of athletic ability development of athletes, to tailor the training content, method, intensity, and density for each athlete more precisely and effectively, so that each athlete can obtain the optimal training effect.

However, most of them only point out which physical fitness indicators have a more significant effect on sports performance and abstractly conclude that coaches should strengthen this aspect more in training, without giving specific quantitative indicators, which is insufficient for our coaches. Modern training programs must move towards quantification and precision, which urgently requires an assessment model of the relationship between physical fitness indicators and athletic performance [3]. With the assessment model, each training has a clear target and provides a scientific basis for the daily training arrangement. In the past, some scholars have used the multiple regression method to establish the assessment model of some sports, but the factors affecting the competitive ability of special sports not only have their functional characteristics but also make up for each other and promote and influence each other, forming a complex dynamic quality system in the human body, so their influence on the final level is often not linear. The method of multiple regression establishes a model that is a linear equation, which necessarily cannot simulate the changing relationship between indicators and performance very accurately [4]. Athletes not only need to have the ability to complete complex and diverse actions with high quality, but also need to realize the seamless integration of action content and musical rhythm. Such competition needs require higher coordination and sensitivity, spatial conversion, and perception capabilities of athletes. Therefore, it is urgent to build an assessment model with a higher degree of simulation. Somatosensory technology is the best embodiment of human action recognition in human-computer interaction. Before the advent of somatosensory technology, human-computer interaction used hardware devices, such as mouse and keyboard, to manipulate the computer, reducing the freedom of human-computer interaction. The advent of somatosensory technology allows people to use their body movements very directly to manipulate the surrounding settings or to interact with the environment around them. Somatosensory technology allows people to get rid of the control of hardware facilities such as keyboards and mice and interact with computers more directly and freely, allowing people to

integrate more realistically into the environment of human-computer interaction [5].

However, special theoretical research mainly concentrated in a few major areas, and there are still many important areas that have not been covered. Theoretical research lags a little behind compared with practice, which is bound to become an obstacle for the continued development of aerobics. Therefore, it is necessary to fill the gaps in these areas. The establishment of a quantitative assessment system for aerobics special movement techniques can scientifically monitor and evaluate the development level of each aerobic athlete's competitive athletic ability. Thus, there is a unified standard for the assessment of an athlete's technical level. Given the lack of this study, this is the research direction of this study and the basis for the selection of the topic. The physical quality of aerobic athletes is an important basis for their special athletic ability, and the prediction model of athletes' special performance can be accurately diagnosed and evaluated using physical form quality index as the independent variable and special performance as the dependent variable, to clarify the focus and target of training content and improve the scientific degree of aerobic training. The initial selection is often made by young children who have not received professional training in aerobics and obviously cannot be selected by examining their aerobic skills. With an assessment model, it is possible to predict their future development potential by testing several physical fitness indicators.

2. Current Status of Research

Memory is ubiquitous in people's lives as a process that enables the storage of experience; for example, the acquisition of life experience and how to learn to use household appliances are all things that memory supports you to achieve [6]. In sports, memory is associated with the acquisition of motor skills. Thus, memory plays an important role in human survival and development. Situational memory is an individual's memory for the experience of an event [7]. This study found that situational memory decreases with age. Situational memory includes memory for events, as well as memory for logical time. Procedural memory is purposeful, practicing motor-operational skills by achieving some purposes, such as skill memory [8]. The quantitative measurement of indicators is ensured to facilitate the quantitative research of specific physical fitness evaluation and training. In the continuous human motion, the energy change during human motion is taken as the idea, and the continuous motion segmentation model of the human body composed of the kinetic energy between joint frames and the potential energy difference between joint frames is established. Memory is initially consciously executed, and after expertise, motor skills become unconscious. At the automatic stage, it is similar to implicit memory, but cannot be totally segregated, with the exception that this component is much reduced. Memory for movement and what can be described as situational memory related to experiential events, and procedural memory related to skill acquisition, are not identical. Action memory can be

enhanced by conscious engagement or some descriptive verbal text, or it can be improved by operant learning; it does not exist singularly but is a complex memory process combining multiple attributes [9]. Nowadays, it is studied as a separate field, and its separate division will also help in future research and applications [10].

Traditional methods usually require the manual design of features and their extraction, followed using various machine-learning algorithms to model the proposed features. Sufficient prior knowledge is also required to support the modeling process to achieve a high action recognition rate [11]. Depending on the type of features extracted, traditional-based action recognition methods can be further classified as human geometry-based methods, motion information-based methods, and spatiotemporal interest point-based methods [12]. Deep learning-based action recognition is an end-to-end approach that automatically learns relevant features directly from the original RGB video sequences and uses them for action classification and is divided into dual-stream convolutional neural network-based action recognition methods, 3D convolutional neural network-based action recognition methods, and long- and short-term memory network-based action recognition methods depending on the network structure [13]. At present, there is no clear definition of the concept of human action in the industry, and the hierarchy of actions cannot be accurately classified. Which stages a human action can be decomposed into, which gestures should be included in each stage, and how to determine the start and end time of each gesture are urgent problems to be solved in the future [14].

In competitive aerobics, the flexibility of the joints of the whole body has its specific role in each movement, and anyone movement that is not in place will slow down the process of the full set of movements, making it difficult for athletes to play the technical level. Good flexibility quality can increase the amplitude of the action, make the action more stretching, posture shape more beautiful, and more artistic expression is the basic guarantee of high quality to complete the action. At the same time, good flexibility quality can also effectively prevent sports injuries and extend the life of the sport. Therefore, in the training of competitive aerobics flexibility quality has been an important training content.

3. Analysis of Deep Learning Recognition and Evaluation of Aerobic Movements

3.1. Deep Learning Action Recognition Algorithm. Based on the construction of human action recognition features, the design uses classifiers to recognize and classify human actions based on the features [15]. In recent years, due to the development of deep learning technology, it has many advantages in the field of human action recognition that traditional methods do not have, so it has been widely used, and currently, the field of human action recognition mainly uses two kinds of convolutional neural networks and recurrent neural networks. LSTM neural network solves the gradient disappearance problem of recurrent neural networks and can remember more long-term states. Using a

classifier for human action recognition classification is essential to classify the serialized data that have been constructed. In human action recognition, each frame in a sequence corresponds to a moment [16]. To complete the recognition of an action depends not only on the data in the present moment, but also has a strong relationship with the data before this moment, and the whole sequence needs to be analyzed to make a judgment. The emergence of recurrent neural networks solves the serialization problem that traditional neural networks cannot handle. In recurrent neural networks, the output depends not only on the current input but also on the output of the previous layer, and the combination of both determines the final output. The frame interval between the two energy values must be greater than 30 frames, so that the acquisition is completed in sequence. In the end, the frames corresponding to the obtained energy values are arranged from small to large, which is the center of the ten actions obtained by the energy method. There is signaling between the nodes of the hidden layer in a recurrent neural network, and the output of the hidden layer is transformed into a part of the next input; i.e., the output contains the current input and the last output. Having such a special structure allows the recurrent neural network to have the ability to handle sequential data well. Figure 1 shows a diagram of the recurrent neural network structure.

The input gates and the output gates are the control gates: the input gates determine the amount of information to be transferred to the memory cell and the output gates determine how much information from the memory cell will be transferred to the current output. The forgetting gate controls the memory unit and is used to decide between remembering and forgetting the memory unit, and the decision is how much data from the previous moment in the memory unit will be transmitted to the present [17]. Compared with the traditional recurrent neural network, LSTM by introducing a gating structure in it, this structure can well solve the gradient disappearance problem of recurrent neural networks during the training process. At the same time, LSTM can store information for a very long time through the memory unit, which is one of the biggest advantages that distinguish it from other networks. However, too long information will make the LSTM suffer from gradient loss, so it can only play its maximum performance in the appropriate range.

The network input is a fusion feature made by combining four features, and after the previous feature extraction and processing, each human action feature becomes 43-dimensional data, and the length of the data varies depending on the number of frames of each action. Before inputting into the network, each group of data used for training or testing is uniformly processed to equal length for ease of processing, and the remaining sequences are zeroed according to the sequence with the longest sequence in each group. In each period, the data input to the network is a 43-dimensional vector. Next, the intermediate values are fed to the output layer through the computation of the LSTM layer, and the output layer used is the softmax function, which judges the activities and outputs the probability of belonging to each action label, and the corresponding probability with the highest value is the final output class of the network.

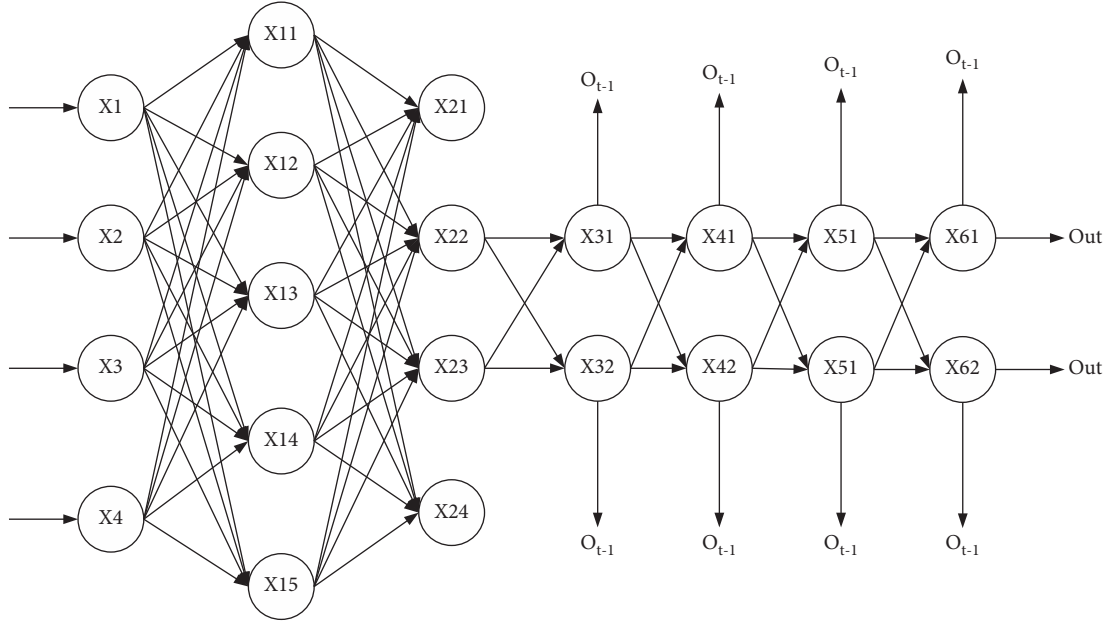


FIGURE 1: Structure of the recurrent neural network.

$$f_t = \sigma(W_f \cdot [h_t, x_t] - b_f),$$

$$a_{tn} = \frac{\exp(\text{ent})}{\sum_{t=1}^T \exp(e_n^t)}. \quad (1)$$

In the process of network training, the parameter setting in the model is an important factor that affects the recognition effect of the model. In this chapter, to facilitate the accurate determination of parameters in the network, the grayscale-processed video frames are input into the network for training. Since the grayscale data and the original data are identical in terms of video length, dataset size, and action type except for the single image pixel data, the grayscale images were chosen to initially determine the training parameters for the network structure. Therefore, in this subsection, the number of samples and iterations of the batch data for the network training process will be determined through grayscale graph experiments. The dataset has a significant impact on the effectiveness of the convolutional neural network. Since the duration of human action is not only related to the action, but also related to the active performer, an action lasts from a few frames to tens of video frames, and even for the same action, the number of video frames that the action lasts is inconsistent due to the different action executors. Then, this will also have a great impact on the recognition effect. In the process of recognition, the processing of a different number of frames will also occupy the different sizes of computing resources. To reduce the use of computational resources, before training the network again, the key region of human motion is first detected using the target detection method, and this region is cut out and added to the network as new data for training. Think of it as another action category in addition to all individual action categories. All these individual action categories and transition action categories are used to train the neural network.

$$c_t = f_t c_{t-1} + i_t \tanh(W_f \cdot [h_t, x_t] + b_f),$$

$$o_t = \sigma(W_o \cdot [h_{t+1}, x_{t-1}] - b_o^2). \quad (2)$$

In human behavioral movements, not all body movements are equally important, and there are often some that are not so important for the recognition of movements. For example, in the human hand waving action, the main change is concentrated in the upper arm part, and the main change part of the kicking action is in the lower limb part; only the action of this one part is needed to discriminate it, and the action of other parts of the body does not have much influence on the determination of the category. Introducing attention to the human action model, so that it looks for key limb and joint points and gives more attention to them, can be more effective for human action recognition. The main principle of the attention mechanism can be understood as that when humans observe a thing, they do not give equivalent attention to the whole thing, but tend to focus most of their attention on a certain part of the observed thing, which helps humans to use their limited attention to quickly obtain valuable information from a huge number of resources [18]. Because of the presence of attention, it makes human processing of image information very efficient and easy. The attention mechanism is to assign different weights to different parts of things so that they play different importance in the final decision, which is generated through the learning process in the network and is constantly updated. In this study, the attention mechanism is combined with Bi-LSTM neural network for human action recognition, compared with the previous LSTM neural network to make a comparison; adding the attention mechanism to the neural network can calculate the weights of the feature vector output in the network at different time points, find out the important features in the whole human action

sequence, and finally improve the accuracy of the whole human recognition network.

$$\begin{aligned} \text{Score}_i &= f(Q_i^2, h_i^2), \\ \alpha_i &= \text{soft min}(\text{Score}_i). \end{aligned} \quad (3)$$

First, some neurons in the hidden layer of the neural network are randomly removed, but the neurons in the input layer, as well as the output layer, remain unchanged. Next, the data of human action features are input on top of this network and the data are propagated forward through the neural network, and eventually, the loss values of the network are back propagated after a part of all the training samples are performed in this process and this part of the samples are updated with the parameters on the retained neurons. Finally, this process is kept repeated to recover the previously removed neurons and again some neurons are randomly selected from the hidden layer for deletion, but to record the parameters of the deleted neurons in the training of some of the samples, as shown in Figure 2. According to the athlete's current level, the coach can predict the athlete's possible performance peak according to the prediction model. The coach can calculate the athlete's current physical fitness and the corresponding sports performance and perform the exercise on the athlete according to the athlete's physical function potential and the coach's experience.

Strong attention differs from soft attention in that strong attention considers that every point in an image may be extended to attention, while strong attention is a powerful stochastic prediction process where the system samples randomly from implicit states rather than decoding using all implicit states, with more emphasis on dynamic changes. The bottom line is that strong attention is non-microscopic attention, and the training process is often done through augmented learning. Since the gradients in the network are computed directly and not estimated by a stochastic process, we choose the soft attention mechanism. A temporal attention model is added to the extracted feature model, using a layer of LSTM network with updated weight values.

3.2. Aerobic Movement Evaluation Design. In the process of testing using the metronome after the experiment, the experimental group was able to judge the correct beat of the music faster than before the experiment, and the students' ability to hear and recognize the rhythm of the music was strengthened, and there was a great improvement in their ability to grasp the rhythm of the music, while the control group used the traditional aerobic teaching method, and the teachers paid more attention to the training of the specification of each technical movement of aerobics and the completeness of the set movements [19]. The music is an indispensable part of aerobics, and in the teaching methods and means for the strengthening of rhythm practice, some teachers still do not pay enough attention to the teaching of the early stage of simply teaching the action, so that students are only in the over and over again skilled, and in the late direct use of music practice, the teacher did not teach the process to effectively establish the relationship between

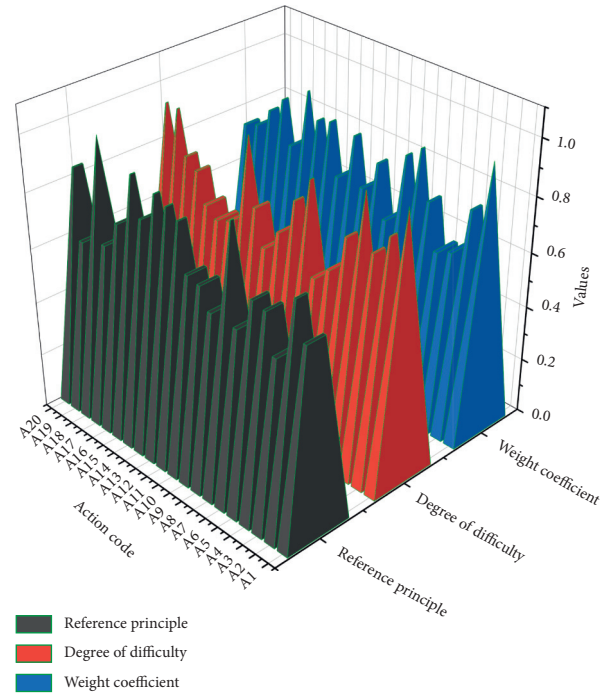


FIGURE 2: Movement indicators for assessing the technical level of specific movements of bodybuilders.

music and sets of movements, so that students lack understanding of music and cannot establish the music. Therefore, the traditional teaching method is not obvious for students to improve their ability to hear and identify music beats, so the experimental group students' music beat reaction time test scores are better than the control group, which has a significant teaching effect.

Physical fitness is the basic athletic ability expressed through the qualities of strength, speed, endurance, coordination, flexibility, agility, etc. It is the core element of competitive ability, and the external form and internal function have an impact on physical fitness. Each project has different requirements for physical fitness. Good physical fitness is the prerequisite for systematic and scientific improvement of performance; it is the necessary guarantee to continuously increase the load and maintain the training efficiency; it is the necessary foundation to improve the competitive state and ensure the level of participation; it is the key to effectively avoid trauma and prolong the training life. The core task of physical training is to improve physical quality and improve external form and strengthen internal functions according to the competitive needs of the project. From the viewpoint of competition time, the set time of aerobics is 80 (± 5) seconds, in which athletes need to carry out high-intensity exercise without interval and make endurance adjustments through the action of main content; from the viewpoint of energy supply mode, glycolysis system is the main source of energy of aerobics; and from the viewpoint of action technology, aerobic technology is diverse, and the mastery and play of technology is the key element that dominates the result of participation. In summary, aerobics is a skill-driven sport with short

duration, high load intensity, and no intervals, as shown in Table 1.

The external shape consists of length, circumference, density, etc., while internal components are mainly muscle, fat, and body fluid components. The morphological characteristics of athletes vary and differ at different ages. The development of body shape is to some extent the result of quantitative changes in training, which is conducive to the improvement of function and quality, and its changes have an impact on the training effect. For the aerobic program, which is dominated by the judges' subjective judgments, good morphological indicators can enhance artistic expression. To better summarize the morphological characteristics of aerobic athletes, measurements were taken on the athletes participating in the study.

The average height of the participating athletes was 176 cm, with no athletes too tall and no athletes too short. Complexity, variety, and innovation are among the requirements for judging all movement content in the rules of aerobics. Athletes not only need to have the ability to complete complex and varied movements with high quality but also need to have the ability to achieve seamless integration of movement content and music beat, which requires high coordination and agility as well as spatial transformation and perception ability. Good physical fitness can not only lead to the improvement of skills, but also a necessary condition to complete the competition goals and improve the competition performance; excellent strength quality is the guarantee for the realization of the stable improvement of difficult skills. Athletes who are too tall have certain disadvantages in agility and coordination, and they have difficulty in completing large and varied movements, which is not conducive to training and competition; while athletes who are too short have the innate advantage of better agility and coordination, those who are short have a certain disadvantage in artistic aesthetics due to their small range and poor expressiveness in the expression of movement content. Therefore, considering the needs of competition and movement expression, athletes of moderate height are more conducive to aerobic training and competition (Figure 3).

There are many special fitness indicators and complex structures, and different indicators have different effects on the level of special fitness of high-level aerobic athletes in China, and the complexity and operation methods vary in the actual test. Which indicators can reflect the current situation of athletes' special fitness in a comprehensive and representative manner, which indicators are more convenient to obtain data, and which indicators are more conducive to the development of special fitness of aerobic athletes are all indicators of the key elements that need to be considered comprehensively in the selection of indicators. Therefore, before constructing the index system, the principles that need to be followed for index selection should be determined first to guide the selection of the indexes. Firstly, we should consider the representativeness of the indicators, and the indicators with high representativeness can reflect the competition demand more intuitively. This will inevitably be affected by the scale of the competition and the level of the participating players. It is very meaningful to establish

a set of evaluation systems for athletes' special sports competence in these sports.

The availability of indicator data is a prerequisite for the scientific conduct of the evaluation.

It is required that the selected or designed indicators can obtain the corresponding data in practical application, avoiding the indicators for which data are not available or accurate data cannot be obtained and ensuring the quantitative measurement of indicators for the quantitative research of special physical performance evaluation and training. In human continuous action with the idea of energy changes during human movement, a human continuous action segmentation model consisting of kinetic energy between joint frames and the potential energy difference between joint frames is established [20]. After visualizing the recognition effect, the recognition results of the single-channel network and the two-speed channel network have the same trend, the difference is that the action recognition changes more smoothly when recognizing the video action after using the two-speed channel, and the overall recognition probability is slightly smaller than the recognition result of the single-channel network, as shown in Figure 4.

To make the evaluation process more objective and accurate, it is necessary to develop a unified evaluation standard that can provide a more intuitive and clear understanding of the differences in the special physical abilities of high-level aerobic athletes. To ensure the objectivity of the evaluation criteria, 16 high-level aerobic athletes were selected to test the index system, and all the athletes selected had the title of the athlete with the rank of general and had won the top 3 places in the authoritative Chinese aerobic events in recent years to ensure the representativeness of the test sample. The data of the athletes were obtained for each index for the follow-up study. This is obviously not enough for our coaches. Modern training programs should be developed in the direction of quantification and accuracy, which urgently needs to establish an evaluation model of the relationship between physical fitness indicators and sports performance.

4. Analysis of Results

4.1. Deep Learning Action Recognition Results. According to the energy model established earlier for the segmentation of continuous action recognition, the energy difference before and after all frames in human continuous action is calculated. Figure 5 is a schematic diagram of the energy of a continuous action obtained after the calculation, the energy between the different two frames has a large difference, and the energy method is based on this for the segmentation of continuous actions. Based on the overall idea that the energy value between two frames when the action occurs is higher than that when the transition action occurs, the segmentation of the continuous action is performed according to the energy value. All the frames of this continuous action are sorted according to the energy value, and the ten frames with the largest energy values are taken to correspond to the ten actions in the continuous action. To avoid that, some individual actions have higher energy values in general, which

TABLE 1: Body morphometry.

Height	Weight	Leg length	Shoulder width	Chest circumference	Waistline	Hip circumference
Average value	70.3	75.6	41.7	72.5	53.4	44.5
Standard deviation	55.8	82.6	62.7	92	58.2	90.7
Minimum	91.5	60	76.8	58	50.6	50.4
Maximum value	41	42.6	88.9	71.6	94	76.5

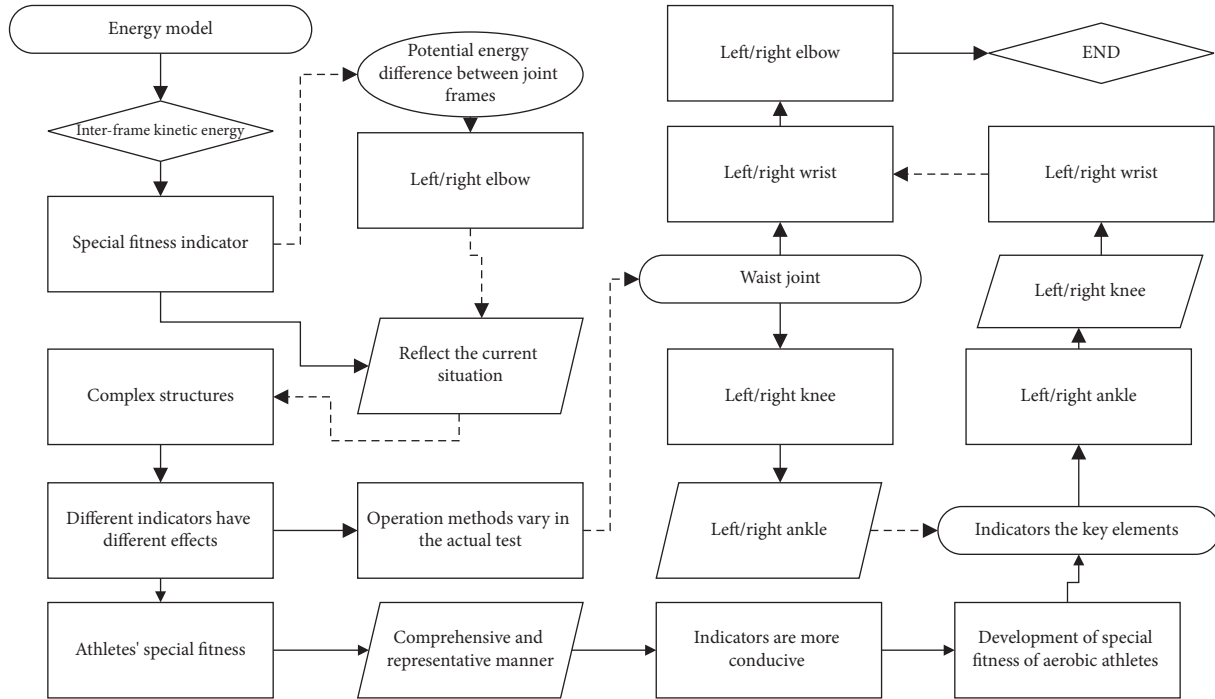


FIGURE 3: Energy model node composition.

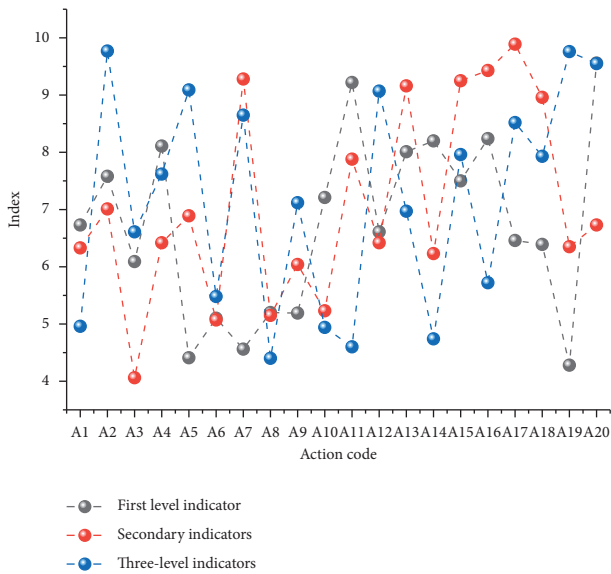


FIGURE 4: Evaluation indicator system and weighting factors.

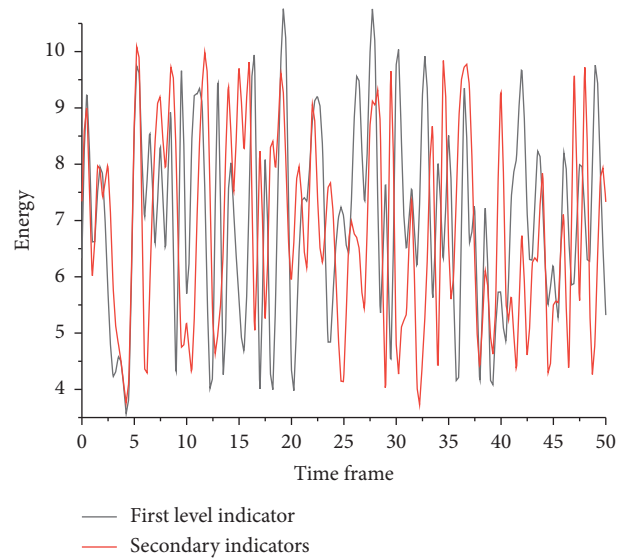


FIGURE 5: Schematic of continuous action energy.

affects the acquisition, and the interval between two energy value frames must be greater than 30 frames, so that the acquisition is completed in turn. Eventually, the frames corresponding to the obtained energy values are arranged

from smallest to largest, which is the center of the ten actions completed by the segmentation obtained by the energy method. The arrows point to the locations of the centers of the actions segmented using the energy method, and the

lower line segment is the correct segmentation interval for successive actions. Some of them were effectively segmented, but some of them could not be effectively identified. This is mainly because jitter occurs from time to time during the acquisition of a human continuous action data point, which can cause a dramatic change in the energy value between a certain two frames, causing some degree of interference to the energy method segmentation.

For a human continuous action sequence, the sliding window combined with a neural network classifier has the theme idea of identifying the segmentation while sliding. The neural network classifier is needed in the segmentation process of the human continuous action, and for the training set of the neural network classifier in addition to using the manually segmented individual action sequences in continuous action, there should be transitional action sequences between individual actions and the first and last interspersed action sequences as another action category in addition to all individual action categories. A complex dynamic quality system is formed in the human body, so their influence on the final level is often not linear, and the model established by the multiple regression method is a linear equation. These all individual action categories, as well as the transition action categories, are used to train the neural network. Finally, the segmentation of sequential action sequences of the human body is completed using the trained neural network. In conjunction with a sliding window, the action data are collected at a fixed window size in certain steps, and the neural network classifier completes the recognition segmentation of the actions in the window at that moment and then continues to the next moment window, and when the results of the action categories recognized in the first two moments do not agree, it is determined that the boundary of that action category is reached. When the recognition segmentation of the whole human continuous action is completed, the segmentation result of the continuous action is finally determined after using the screening mechanism to filter and evaluate the segmented actions in the whole human continuous action, as shown in Figure 6.

In the segmentation model of human continuous action recognition based on sliding window combined with the neural network classifier of this study, four of the continuous action sliding window recognition results obtained are shown in Figure 6 below, and a total of ten action classes and one transition action class are recognized. In the recognition experiments of ten actions based on sliding window combined with the neural network classifier of this study for all continuous actions, different step lengths may have an impact on the segmentation recognition of continuous actions, combined with the continuous action database used in this study, the minimum action sequence frame number is below 10 frames, and using a larger step length will cut or jump a certain complete action. Somatosensory technology allows people to get rid of the control of keyboard and mouse and other hardware facilities, and interact with computers more directly and freely, allowing people to integrate into the human-computer interaction environment more realistically. The overall effect of the sliding window on the recognition rate is that the recognition rate decreases as the

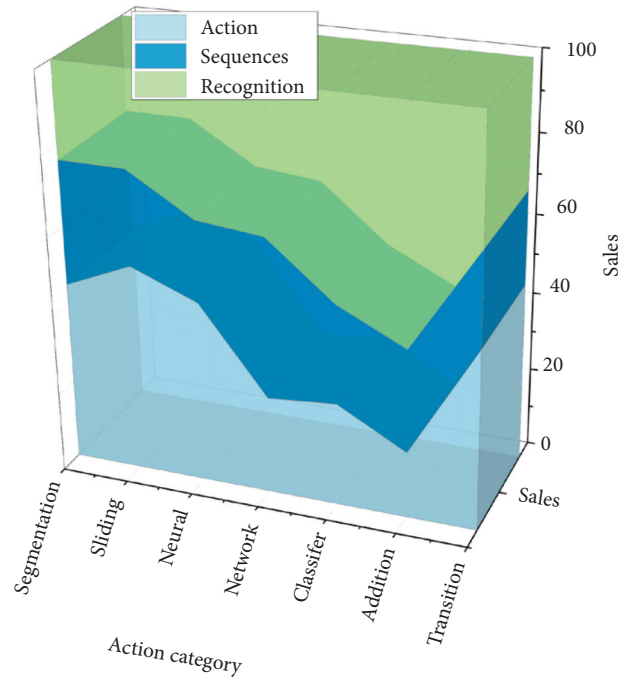


FIGURE 6: Continuous motion sliding recognition results.

window increases, mainly because the length of each action sequence in a continuous action is mostly concentrated around 30 frames, and using a larger window tends to include transitions outside the boundary, resulting in a lower recognition rate. The maximum recognition rate is obtained at 15 frames mainly because when recognizing an action, the important features tend to be concentrated in a few frames, and the action can be recognized and classified by a few frames.

In this database segmentation problem, two continuous action segmentation methods are given, the energy-based segmentation method and the sliding window-based method combined with the neural network classifier in this study. In the energy-based continuous action segmentation method, an energy model consisting of the kinetic energy between joint frames and the potential energy difference between joint frames is established, and the energy at the time of action is higher than the energy at the time of transition action as the general idea for the segmentation of continuous actions, and the obtained action centers are given time windows of different sizes for experimental simulation. Then, the sliding window combined with the neural network classification built in this study is used to identify the segmentation of human continuous actions, and the initial boundary points obtained are evaluated using a screening mechanism to select the final action boundary in the experimental simulation.

4.2. Results of the Evaluation of Aerobic Movements. The completion of difficult movements and the level of physical training is not a simple superposition of sports qualities, but the result of the comprehensive role of the qualities. Each quality not only has its own functional properties but there is

also a complicated interaction between them that not only makes them up but also influences and restricts them. In the human body, a complicated dynamic quality system is formed that affects the overall functional effect. The final score of the difficult action technique is closely related to the structure of physical qualities. The question of what level of each quality should be reached to achieve the ideal score for difficult movements is related to whether the athletes' physical quality potential can be fully utilized and brought into play. Therefore, it has been widely used. Currently, convolutional neural networks and recurrent neural networks are mainly used in the field of human action recognition. Therefore, the development of evaluation criteria for the final score of difficult movement techniques and the level of physical quality development plays a crucial role in the development of training plans to improve the final score of difficult movements. The difference in the difficulty score itself gives a different value to the difficult movements. The higher the difficulty score, the higher the technical and physical requirements of the athlete to complete the movement. If the difficulty factor is too low, it will not be possible to open the gap between the technical levels of the movements of individual athletes. Therefore, we should try to choose the movements with a higher difficulty coefficient as the evaluation index, as shown in Figure 7.

Usually, the development of a professional-level aerobic athlete is long-term and requires the selection of potential recruits starting at a young age. In early selection, we can test the six basic quality indicators of the prospect, which can predict the potential that the prospect has in the sport of aerobics through the prediction model, and combined with the coaches' own accumulated experience in selection, we can make the early selection more justifiable. LSTM can store information for a long time through the memory unit, which is the biggest advantage that distinguishes it from other networks. However, too long information will cause the LSTM to lose the gradient, so it can exert its maximum performance only in an appropriate range. Based on the existing level of the athlete, the peak performance that the athlete may reach can be predicted by the coach based on the prediction model to calculate the existing physical quality of the athlete and the corresponding athletic performance and make a judgment on the athlete based on the potential of the athlete's physical function and the coach's experience to accurately predict what height the athletic performance of the athlete can reach after a certain period of training. The coaches can measure the changes in the athletes' physical quality indicators and sports performance regularly, compare the measured values with the expected values calculated by theory, and check the sports training effects in stages to make timely adjustments to the training plan so that the whole training process is under control. Coaches should determine the content of training according to the two situations of athletes: for athletes with unbalanced physical quality development, to effectively improve the overall level of physical quality, strengthening the weak links in their physical quality should become the focus of training. For athletes with a balanced level of physical quality development, the content of sports training should be arranged

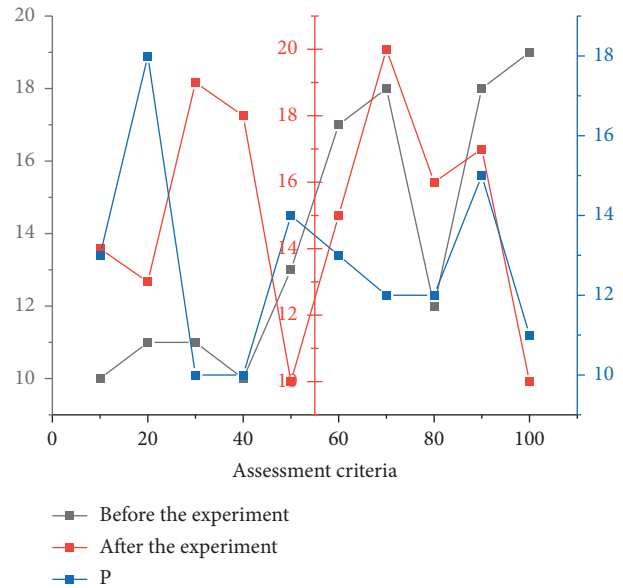


FIGURE 7: Intragroup comparison of the 8 items of aerobic set performance of the experimental group before and after the experiment.

according to the principle of priority development of physical quality.

5. Conclusion

Based on the improved C3D model, a deep convolutional neural network extracts per-frame-based features in the input video. After feature extraction, these features are fed into a temporal attention subnetwork. The temporal attention subnetwork assigns the relevance weight of the features in the recurrent neural network based on the relevance of each feature to the action topic in a positive correlation. This weight is constantly updated during the training iterations of the network during the activity recognition process. The product of superimposed weights and features is used in the temporal attention model to derive the action state information at the current moment, and finally, the decision classification is added to obtain detection results based on the motion state information. This helps humans use limited attention to quickly obtain valuable information from massive resources. Because of the existence of attention, the processing of image information by human beings is very efficient and convenient. In this study, 12 difficult action indicators were selected from 307 movements to reflect the movement skill level of aerobic athletes, and the evaluation table of special movement skill level was developed using the weighting method, and the reliability test results showed that the established evaluation table can accurately reflect the movement skill level of aerobic athletes. The prediction model was established using the method of the neural network, and the reliability and error of this model were analyzed. The error test showed that the evaluation model of body form and quality of aerobic athletes established by the neural network can predict aerobic sports performance and assess the development level of physical quality more accurately.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

All the authors do not have any possible conflicts of interest.

Acknowledgments

This study was supported without any funding.

References

- [1] S. Hwangbo, R. Al, X. Chen, and G. Sin, "Integrated model for understanding N₂O emissions from wastewater treatment plants: a deep learning approach," *Environmental Science & Technology*, vol. 55, no. 3, pp. 2143–2151, 2021.
- [2] P. Pareek and A. Thakkar, "A survey on video-based Human Action Recognition: recent updates, datasets, challenges, and applications," *Artificial Intelligence Review*, vol. 54, no. 3, pp. 2259–2322, 2021.
- [3] Z. Wang, L. Zhang, M. Zhao et al., "Deep neural networks offer morphologic classification and diagnosis of bacterial vaginosis," *Journal of Clinical Microbiology*, vol. 59, no. 2, pp. e02236–20, 2020.
- [4] Q. Vanhaelen, Y.-C. Lin, and A. Zhavoronkov, "The advent of generative chemistry," *ACS Medicinal Chemistry Letters*, vol. 11, no. 8, pp. 1496–1505, 2020.
- [5] D. Dave, D. J. DeSalvo, B. Haridas et al., "Feature-based machine learning model for real-time hypoglycemia prediction," *Journal of Diabetes Science and Technology*, vol. 15, no. 4, pp. 842–855, 2021.
- [6] T. Cai, J. Long, J. Kuang, F. You, T. Zou, and L. Wu, "Applying machine learning methods to develop a successful aging maintenance prediction model based on physical fitness tests," *Geriatrics and Gerontology International*, vol. 20, no. 6, pp. 637–642, 2020.
- [7] K. Candoğan, E. G. Altuntas, and N. İğci, "Authentication and quality assessment of meat products by fourier-transform infrared (FTIR) spectroscopy," *Food Engineering Reviews*, vol. 13, no. 1, pp. 66–91, 2021.
- [8] J. Xu, H. Zhang, J. Zheng, P. Dovoedo, and Y. Yin, "eCAMI: simultaneous classification and motif identification for enzyme annotation," *Bioinformatics*, vol. 36, no. 7, pp. 2068–2075, 2020.
- [9] E. Gato, I. P. Constanoso, A. Candela et al., "An improved matrix-assisted laser desorption ionization–time of flight mass spectrometry data analysis pipeline for the identification of carbapenemase-producing *Klebsiella pneumoniae*," *Journal of Clinical Microbiology*, vol. 59, no. 7, pp. e00800–21, 2021.
- [10] P. D. Loprinzi, J. Harper, and T. Ikuta, "The effects of aerobic exercise on corpus callosum integrity: systematic review," *The Physician and Sportsmedicine*, vol. 48, no. 4, pp. 400–406, 2020.
- [11] K. McCullough, T. Williams, K. Mingle, P. Jamshidi, and J. Lauterbach, "High-throughput experimentation meets artificial intelligence: a new pathway to catalyst discovery," *Physical Chemistry Chemical Physics*, vol. 22, no. 20, Article ID 11174, 2020.
- [12] B. Pardamean, H. Soeparno, A. Budiarto, B. Mahesworo, and J. Baurley, "Quantified self-using consumer wearable device: predicting physical and mental health," *Healthcare informatics research*, vol. 26, no. 2, pp. 83–92, 2020.
- [13] M. Jacobsen, T. A. Dembek, G. Kobbe, P. W. Gaidzik, and L. Heinemann, "Noninvasive continuous monitoring of vital signs with wearables: fit for medical use?" *Journal of diabetes science and technology*, vol. 15, no. 1, pp. 34–43, 2021.
- [14] N. Mischenko, M. Kolokoltsev, E. Romanova, V. Dychko, and Y. Dychko, "Using» Flipped Classroom «pedagogical technology in school physical education," *Journal of Physical Education and Sport*, vol. 20, no. 6, pp. 3504–3511, 2020.
- [15] J. Zhao, F. Pan, F. Pan et al., "Detection of cotton waterlogging stress based on hyperspectral images and convolutional neural network," *International Journal of Agricultural and Biological Engineering*, vol. 14, no. 2, pp. 167–174, 2021.
- [16] P. Gupta, A. Thatipelli, A. Aggarwal et al., "Quo vadis, skeleton action recognition?" *International Journal of Computer Vision*, vol. 129, no. 7, pp. 2097–2112, 2021.
- [17] O. Morin, M. Vallières, S. Braunstein et al., "An artificial intelligence framework integrating longitudinal electronic health records with real-world data enables continuous pancreatic cancer prognostication," *Naturaliste Canadien*, vol. 2, no. 7, pp. 709–722, 2021.
- [18] W. R. Black, C. A. DiCesare, S. Thomas et al., "Preliminary evidence for the fibromyalgia integrative training program (FIT teens) improving strength and movement biomechanics in juvenile fibromyalgia," *The Clinical Journal of Pain*, vol. 37, no. 1, pp. 51–60, 2021.
- [19] R. Liu, Y. Gu, M. Shen et al., "Predicting postmortem interval based on microbial community sequences and machine learning algorithms," *Environmental Microbiology*, vol. 22, no. 6, pp. 2273–2291, 2020.
- [20] P. Husu, K. Tokola, H. Vähä-Ypyä et al., "Physical activity, sedentary behavior, and time in bed among Finnish adults measured 24/7 by triaxial accelerometry," *Journal for the Measurement of Physical Behaviour*, vol. 4, no. 2, pp. 163–173, 2021.