Research Article

Simulation on Basketball Recognition Based on Mobile Network Communication and Panoramic Vision

Yang Yang

Institute of Physical Education, Huanghuai University, Zhumadian 463000, Henan, China

Correspondence should be addressed to Yang Yang; 20101168@huanghuai.edu.cn

Received 30 June 2022; Revised 30 July 2022; Accepted 9 August 2022; Published 9 September 2022

Academic Editor: Raghavan Dhanasekaran

Copyright © 2022 Yang Yang. 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.

At present, the application of information technology has realized the transformation of people’s production and lifestyle, and it has also promoted the transformation of the sports industry. At present, the application of informatization in the basketball field in the sports industry is increasing. By using the advanced methods and technologies of its information display, this article aims at realizing the optimization of the basketball recognition pattern in the mobile network communication environment to promote the ecological development of the sports industry. It proposes to use panoramic vision to guide the integration of computer network and basketball field, which is helpful to analyze and solve the objectivity problems of single action and system incompatibility in basketball recognition simulation. By referring to the principles and laws of panoramic vision, the identification module and simulation module of the basketball auxiliary training system are constructed and optimized, so as to realize the promotion of basketball and the ecologicalization of teaching. In the research on the basketball recognition simulation system, the three-dimensional skeleton key point coordinate data are input into the ST-GCN network for comparison and testing. Specifically, it includes 2D key point coordinate data and 3D key point coordinate data as input to conduct model research and testing in the self-made basketball dataset. The experimental results show that the recognition rates of 2D and 3D coordinates are 66.64% and 87.69%, respectively. Therefore, it is crucial to use the human panoramic vision algorithm to convert 2D coordinates to 3D coordinates.

1. Introduction

At present, with the awakening of national awareness of physical fitness, all kinds of sports have been promoted and popularized. The basketball equipment is simple, and a wheel-style game on a venue can accommodate 10–20 people to exercise. This sport can exercise all the abilities of the human body in an all-round way, and small players can take advantage of their speed to shuttle through the crowd as a defender. Physically strong players use their physical advantages to perform strong back play, and no matter what their physical conditions are, they can find a suitable position on the court. With the upgrading of mobile devices and the rise of the short video industry, more and more people are sharing their lives on short video platforms. More and more self-media has also contributed rich spiritual and cultural food to people by outputting high-quality short videos. Among them, the short videos of basketball teaching and basketball highlights are imitated and learned by basketball fans. Models such as convolutional neural network-based models play an important role in panoramic views. Convolutional neural has been used in the image field, and after achieving excellent performance in this field, researchers have successively proposed algorithm models that apply it to video recognition. If the basketball technical movements that appear in the video can be recognized, it will be an important application of panoramic view in sports. Such practical applications have great prospects in the cyber basketball community.

With the development of panoramic perspective and artificial intelligence theory, analysis based on various human behavior datasets emerges in an endless stream, and the action recognition technology based on image frames is constantly innovating, making video behavior action
recognition feasible. In basketball technical action videos, technical actions have obvious characteristics. In this kind of video on the short video platform, the key characters in the shot in the sports video are relatively fixed, and the scene is relatively simple, so the technical action classification of basketball technical action videos has certain outstanding advantages. However, there are also many challenges in basketball technical action recognition, mainly including how to effectively utilize consecutive image frames with strong correlation. In addition, there are many complex basketball techniques, and it is difficult to screen out representative basketball techniques. In view of the above problems, this article firstly determines the concept of technical action through the research of sports literature. By studying the practical application of technical movements in professional competitions and the teaching videos of short video platforms from the media authors, the basic elements of basketball technical movements videos are analyzed. And through the literature research of basketball action dataset, the characteristics of the dataset are analyzed, and the basketball technical action dataset is established based on the collection method of badminton and table tennis technical action datasets.

The identification and simulation research of basketball has become an important node to promote the development of the sports industry, and many scholars have carried out research in this area. Among them, Zhu studied a basketball player gesture recognition algorithm based on a multisensor method. He completed the whole process from data acquisition to data processing and model algorithm construction and verification [1]. Castro et al. analyzed the effect of wearing lace-up ankle braces. And they developed a motion program to simulate the intensity of the ground reaction force during a specific basketball vertical jump in a basketball game [2]. Wang proposed a design optimization method for basketball teaching and training system based on motion capture technology. By comparing the simulation results with real training videos of athletes, he displayed training movements and standard movements on the same screen and conducted a comparative analysis of the movements [3]. Pengyu and Wanna research mainly took the characteristic information of basketball in the state of basketball goal as the starting point, and compared and analyzed the detection methods by detecting the target in the environment [4]. Ma proposed and improved the optimization method of basketball skills and teaching mode based on visual action simulation. Through the construction of the basketball action analysis system, teachers can effectively find the mistakes of students’ actions and give guidance [5]. However, their research on basketball recognition simulation only focuses on efficiency improvements. It does not take into account its deep-level characteristics, so the research is only in the theoretical stage and has little practical significance.

Based on mobile network communication technology and panoramic vision technology, it is urgent to carry out identification and simulation research on basketball. Among them, Li et al. conducted an empirical study on the technical analysis and tactical training of college basketball based on the apriori algorithm. Association mining technology can effectively help athletes analyze relevant data, and then analyze and correct wrong actions [6]. Based on learning techniques, Li and Gu used sensors with gesture recognition algorithms to analyze detailed motion capture of sports players [7]. Liu et al. proposed panoramic vision. They combined 3D recognition technology with the selection of basketball techniques and tactics to help improve the level of basketball techniques and tactics [8]. Hao et al. proposed that multitarget tracking combined with target corner features can track different parts of the athlete as different target areas [9]. Shouyuan discussed the recognition and simulation of basketball training behaviors based on virtual reality. He used mobile network communication technology to establish related college basketball education innovation programs [10]. However, the current research on the sports industry in the context of computer network still does not get rid of the definition and thinking based on the traditional basketball field. This article lacks in-depth analysis and discussion on the functionality of panoramic technology.

The innovation of this article is that (1) the definition of the basketball technical action dataset is completed by referring to basketball textbooks and literature about basketball in China. This article refers to the collection method of technical action videos of table tennis, combined with the popularity of the mobile Internet and the rise of short videos, to design a collection process of technical action datasets. After that, this article will complete the preprocessing of the basketball technical action video dataset based on the rules of basketball games, the analysis of the rationality of the action, and the technical action that appears in the professional basketball league to prove the representativeness of the technical action. (2) This article establishes a basketball technical action dataset, including 6 kinds of technical actions; each action has 300 segments, a total of 1800 videos. And it combines object detection to generate low-resolution image input for action recognition with a dual-resolution 3D convolutional neural network architecture. (3) After a series of processing procedures, the original frame images of the basketball technical action video set are generated by a method based on target detection to generate a low-resolution cropped frame dataset. Finally, combined with the 3D-CNN network model, the video human action recognition algorithm architecture of the double-resolution 3D-CNN in the literature is improved, and the effectiveness of the algorithm in this article is verified by experiments on the basketball technical action video set.

2. Basketball Simulation System Based on Mobile Communication and Panoramic Vision

This system is introduced into the video-based basketball auxiliary training through the research on the recognition algorithm. It provides auxiliary guidance for the basketball training teams of major colleges and universities and the training of basketball clubs in China [11]. Through the collection of team and player schedule information, basic
data collection, and data comparative analysis, in personal training, users upload single-player sports videos. By comparing and analyzing the extracted three-dimensional skeleton information and the skeleton information of standard movements, it assists players to train for standard movements. In team games, people upload the complete game video, analyze it through the introduced recognition algorithm, feedback the movements and positions of each player, and provide tactical guidance to the coach. Before the development of the system, it is necessary to conduct a feasibility analysis on the system to be developed. Specifically, it includes economic feasibility, technical feasibility, and operational feasibility. The following is the specific analysis content.

2.1. Demand of Basketball System. After the requirement investigation, the requirement analysis of the system is very important in the whole system design and development process. Requirement analysis is to analyze and organize the user needs in the research, and determine the functions of the system and database design with clear thinking, brief text, and standardized structure documents. This section mainly introduces the requirements of ordinary users and administrator functions. User functional requirements are as follows:

(1) Users can query the schedule information and information details through the system.

(2) Users can view the information of teams and players, and can compare the data of teams and players.

(3) Users can upload sports videos. Through the analysis of the systematic action recognition algorithm, the actions of the players in the video are classified and the positions of the players are analyzed. Display the recognized sports video: by uploading a single-player training video, extracting the 3D skeleton video of the player to assist users in longitudinal comparison analysis.

(4) Users can query historical data according to the filtering conditions and make event predictions based on historical data.

(5) The user interface needs to be simple and beautiful, easy to understand and use.

The user example diagram is shown in Figure 1.

2.2. Basketball System Design. The system design of basketball is economically feasible, compared with the commercial sports video analysis system in China. At present, in basketball games, mainstream commercial systems mainly use Chinese analysis systems Synergy Sports, Shot Tracker Team, Coach’sEye, and other systems, which cost relatively high fees. For the training and use of ordinary college basketball teams, the economic pressure is too great. However, the development of this system is completed on personal computers, and the software used are all free.
software. The data in development come from daily collection, the development cost is low, and the basketball auxiliary training system charges lower after commercial use. It solves the problem of high investment cost in colleges and universities [12]. After the research of a 3D-based action recognition algorithm, this system can improve the recognition rate of athletes and extract the corresponding three-dimensional skeleton of the movement. The longitudinal comparison of the skeleton can effectively assist the coach in training. In terms of operational feasibility, the interface of the system is simple in design, and users do not need to understand the background processing process. For users, the data processing process and calculation difficulty in the background can be ignored. At the same time, the system is deployed on Alibaba Cloud and does not require excessive computer configuration, so it is feasible to operate.

During the overall design phase, all requirement analysis tasks need to be integrated into the system to achieve a common system design plan. In the system design, the logic problems in the setting development process should be minimized, and each module should be designed as a black box so that the relationship between the modules can be better defined. This paper mainly discusses two aspects: system architecture design and general system design. This system uses the network architecture shown in Figure 2.

As can be seen from Figure 2, the system should be deployed on Alibaba Cloud, and the system management platform defines unified data and interface specifications. Users who are added and authenticated by the background administrator can request to view team data and player data, as well as player training data and player ability evaluation through the interface [13]. There is also a system for data collection and storage, data analysis, and data display. The functional structure is shown in Figure 3.

As can be seen from Figure 3, it is mainly divided into the following main modules: schedule management module, team management module, player management module, training management module, data analysis management module, and system maintenance module.

2.3. Schedule Information Management Module. The schedule information management module mainly completes the collection and display of the schedule information. For example, the user can query the game time of the team, the location of the game, the opponent, the scores of both sides of the completed game, and view the specific content of the information according to the recent information list.

2.4. Team Information Management Module. The team information management module completes the collection and comparative analysis of the team's basic information. For example, a user can search for a certain team according to the region and view the data information of the team, the
team’s lineup, and the players with the title of “Data King” in the team. At the same time, people can view the team’s ranking according to the team’s game data. People can also filter two teams for comparative analysis and view the radar charts displayed by data such as average points per game and average rebounds.

2.5. Player Information Management Module. The player information management module mainly completes the basic information collection and comparative analysis of players. For example, a user can look up the details of a player by nationality, location, and team name. It includes specific information such as the player’s name, team, birthday, height, weight, experience, and game location, as well as league comparisons and scoring hotspots.

2.6. System Maintenance Module. The system maintenance module mainly completes the user management and basic information management of the system, including schedule information, basic information of players and teams, and training information.

2.7. Detailed System Design

2.7.1. Training Information Management. The training information management module is divided into three sub-modules: basic movements, training plans, and physical fitness reports. The user logs in to the system through the user name and password, and the system automatically determines the user authority. If users have permission to upload videos, they can upload videos, and the system will display the 3D skeleton map and the labeled action recognition videos through algorithms, analyze the game data of the player or team, and specify training [14]. Physical fitness report: It is aimed at individual differences of players, such as speed, endurance, bounce, and strength, combined with the advantages and disadvantages of technical level, injury history, recommend players’ training programs, and track players’ skills and ability comparisons. The process is shown in Figure 4.

2.7.2. Data Analysis Management. This module is divided into two submodules: historical match analysis and match prediction. Historical game analysis is to analyze the records of teams and players in historical games to form game data analysis graphs of players and teams. The match prediction is the prediction of the match results based on the previous match records of the opposing teams, as shown in Figure 5.

2.7.3. Database Structure Design. According to the design of the logic requirements, the physical structure design of the database of the basketball auxiliary training system in this paper is completed, and the training information submodule will be elaborated [15]. The video information table is used to store video information uploaded by users, including video id, name, video link, video status, and upload time. Details are stored in in Table 1.

The specific three-dimensional skeleton information is stored in the skeleton table, including the extracted skeleton information of shooting, layup, in situ dribble, running dribble, blocking shot, and running without the ball (see Table 2 for details).
Specific category information is stored in the action category table, including shooting, layup, dribbling, running, dribbling, blocking, and running without the ball (see Table 3 for details).

Table 3: Action category table.

| Field ID  | Type of data | Name               |
|-----------|--------------|--------------------|
| Type of data | Int (15)   | Varchar (260)   |
| Field description | Action ID   | Action name     |
| Is not empty | NOT NULL | NOT NULL          |

2.8. Neural Network Model. The term “neural network” first came from biology, the reason why human beings can become advanced creatures. It is because the human brain has tens of billions of neurons and has superb learning, reasoning, and logical cognitive abilities. The artificial neural network is a computational model inspired by the neurons of the human brain and designed and developed. The input $x_i$ is multiplied by the weight $w_i$ of the corresponding channel and summed, and then multiplied by the activation function $f$ to obtain the output $y$ of the neuron, whose expression is as follows:

$$y = f \left( \sum_{i=1}^{n} w_i x_i + b \right)$$

In the process of training the neural network, since the weights can change continuously, the neurons can be adjusted continuously by changing the weight parameters, so that the output can achieve the optimal effect [16].
addition, in order to effectively solve the problems of nonlinear division and insufficient expressive ability of network models, activation functions are introduced into neural networks. The function is a preselected nonlinear function, and one can choose different activation functions to suit different application scenarios. The following are several commonly used activation functions.

2.8.1. Sigmoid Function. Its formula is as follows:

\[
\sigma(x) = \frac{1}{1 + e^{-x}}.
\]  

The schematic diagram of its function curve is shown in Figure 6.

It can be seen from Figure 6 that the sigmoid function will have gradient saturation. When the value of \(x\) becomes larger or smaller, the value of the sigmoid function tends to be stable and equal to 0, and hardly changes.

2.8.2. Backpropagation. In order to minimize the total error of the network, the error needs to be back-propagated layer by layer to the input layer. In the process of backpropagation, the gradient descent method can be used to reduce this error until the error tends to be minimized. According to the propagation algorithm of the neural network, it can be known that

\[
y^k_j = f\left(\sum_{i=1}^{d} w_i^k x_i^k + b^k\right),
\]

where \(f\) represents the activation function of the network, \(x\) represents the input value, \(o\) represents the weight parameter, and \(b\) represents the bias parameter of the network. Therefore, the mean squared error of the network model is

\[
E_k = \frac{1}{2} \sum_{j=1}^{l} \left( y_j^k - y_j^{\hat{k}} \right)^2.
\]

Figure 6: Sigmoid function diagram. (a) Schematic diagram of Sigmoid function curve and (b) schematic diagram of Sigmoid function gradient curve.

Figure 7: YOLOv1 detection process.
In backpropagation, the gradient descent method can be used to make the network model adjust the parameters in the direction of the gradient that minimizes the error to minimize the error [17]. Assuming that the learning rate is \( \eta \), the derivation of the mean square error can be obtained:

\[
\Delta w_{hj} = -\eta \frac{\partial E_k}{\partial w_{hj}}
\]  

(5)

2.8.3. Target Detection Algorithm. In the field of target detection, the YOLOv1 algorithm takes the lead in realizing the end-to-end learning of the network, which greatly improves the detection accuracy and promotes the development of target detection algorithms. It realizes the end-to-end learning of the network and has a high detection accuracy. The detection process is shown in Figure 7.

As can be seen from Figure 7, first, the algorithm uses the dataset to train the network model and uses the trained model to detect the target. Second, the algorithm divides the input image to be detected into SSX grids, and each grid network predicts the probability of B bounding boxes and classes. Finally, the algorithm obtains the value by calculating the probability of the bounding box and the class; if the value is greater than the threshold, it is the final detection result. The label save format is

\[
\text{label}_1 = [(x_1, y_1, x_2, y_2, P(\text{truck}) = 1), \]

\[*, (x_3, y_3, x_4, y_4, P(\text{basketball}) = 1)].
\]  

(6)

When the sample image is input to the YOLOv1 model, it contains five parameters, which are \((x, y, h, w, \text{score})\), where \((x, y)\) represents the center coordinates of the bounding box, \(h\) and \(w\) represent the height and width of the candidate box, respectively, and score represents the confidence [18]; the formula is as follows:

\[
\text{score} = \Pr(\text{Object}) \ast \text{IOU}_{\text{pred}}^{\text{truth}}.
\]  

(7)

where \(\Pr(\text{Object})\) represents whether there is a detection target center point in the grid, and IOU represents the intersection ratio between the candidate frame \(A\) and the area \(B\) where the label target is located. The formula is as follows:

\[
\text{IOU}_{\text{truth}} = \frac{\text{area}(A \cap B)}{\text{area}(A \cup B)}
\]  

(8)

To obtain a class confidence score, the class probability can be multiplied by each class confidence with the following formula:

\[
\Pr(\text{Class}(\text{Object})) \ast \text{IOU}_{\text{pred}}^{\text{truth}} = \Pr(\text{Class}) \ast \text{IOU}_{\text{pred}}^{\text{truth}}.
\]  

(9)

When the information of each bounding box is obtained, the error \(E_{\text{all}}\) between the predicted value of the YOLOv1 model and the true value of the label is calculated. The error includes classification error \(E_{\text{class}}\), confidence error \(E_{\text{IOU}}\), and coordinate error \(E_{\text{coord}}\), and the specific formula is as follows:

\[
E_{\text{all}} = E_{\text{class}} + E_{\text{IOU}} + E_{\text{coord}}.
\]  

(10)

For the classification error \(E_{\text{class}}\), its calculation formula is as follows:

\[
E_{\text{class}} = \sum_{i=0}^{S^2} \sum_{n=0}^{C} (p_i(n) - \hat{p}_i(n)),
\]  

(11)

where \(p_i(n)\) represents the probability of the predicted category of the rectangular box and \(\hat{p}_i(n)\) represents the actual category probability. For the confidence error \(E_{\text{IOU}}\), its calculation formula is as follows:

\[
E_{\text{IOU}} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{\text{obj}} \left(\text{score}_i - \text{score}_i^{\text{pred}}\right)^2.
\]  

(12)

Among them, the confidence prediction value score and the true value score are, respectively,

\[
\text{score}_i = \text{IOU}_{\text{pred}}^{\text{truth}},
\]  

\[
\text{score}_i^{\text{pred}} = \begin{cases} 
1, & (I_{ij}^{\text{obj}} = 1), \\
0, & (I_{ij}^{\text{noobj}} = 1).
\end{cases}
\]  

(13)

For the coordinate error \(E_{\text{coord}}\), its calculation formula is as follows:

\[
E_{\text{coord}} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij}^{\text{obj}} \left[(x_i - x_i^{\text{truth}}) + (y_i - y_i^{\text{truth}})^2\right].
\]  

(14)

For the total error \(E_{\text{all}}\), the YOLOv1 model updates the model parameters through backpropagation and gradient descent, so that the loss function of the model converges to

| Layer | Support | Chan. Map | Stride | Example | Search | Channels |
|-------|---------|-----------|--------|---------|--------|----------|
| Conv1 | 12×12   | 94×4      | 2      | 125×125 | 255×255 | ×5       |
| Pool1 | 4×4     | 255×45    | 2      | 58×58   | 123×123 | ×96      |
| Conv2 | 4×4     | 26×26     | 1      | 23×23   | 61×61   | ×94      |
| Pool2 | 5×5     | 382×193   | 2      | 17×17   | 28×28   | ×255     |
| Conv3 | 5×5     | 10×10     | 1      | 15×15   | 26×26   | ×194     |
| Conv4 | 4×4     | 382×255   | 1      | 10×10   | 24×24   | ×194     |
| Conv5 | 5×5     | 254×194   | 2      | 4×4     | 22×22   | ×126     |

Table 4: Siamese network structure.
2.9. Fully Convolutional Siamese Network Model. The target tracking algorithm based on the fully convolutional Siamese network regards the tracking problem as a similarity learning problem, and it proposes a similarity measure function $f(z, x)$. That is, perform the same $p$ transformation on the two inputs $z$ and $x$, and then, pass the result to the function $g$: \[
    f(z, x) = g(\phi(z), \phi(x)).
\] (15)

The network structure of the target tracking algorithm (SiameseFC) of the fully convolutional Siamese network is shown in Table 4.

As can be seen from Table 4, the network structure of SiameseFC is mainly composed of convolutional layers and association layers, and there is no fully connected layer. For associative layers, SiameseFC uses a max-associative layer, except for the fifth convolutional layer, and each convolutional layer is followed by a nonlinear activation layer [20]. Furthermore, batch normalization is used before each ReLu layer when training the SiameseFC network to reduce the risk of overfitting.

The advantage of the SiameseFC tracking algorithm is that the size of the search image can be different from the size of the template image, which allows a larger search image as input and is able to calculate the similarity between the two images. Among them, the similarity function of the algorithm adopts cross-correlation, and its formula is as follows:

\[
    f(z, x) = \phi(z) * \phi(x) + b_1.
\] (16)

In the process of training the SiameseFC tracking algorithm, it adopts the stochastic gradient descent method, and its formula is as follows:

\[
    \arg \min_{\theta \in \Theta} E_{(z,x,y)} L(y, f(z, x; \theta)).
\] (17)

SiameseFC uses the logistic regression loss function, and the specific formula is as follows:

\[
    l(y, v) = \log(1 + \exp(-yv)),
\]

\[
    L(y, v) = \frac{1}{|D|} \sum_{u \in D} l(y[u], v[u]).
\] (18)

It can be seen from the above that the recognition method based on video is lower than the recognition method based on key points of the human body. However, the traditional method of key point of the human body needs to be collected by equipment, which is difficult to be applied in basketball games and training. In this regard, combined with basketball game data, this article mainly focuses on the convolutional neural network structure. Combined with the
panoramic vision algorithm, the basketball motion dataset is used to train the network model, so that the obtained network model can be used for action recognition and compared with other recognition algorithms in experiments.

3. Basketball-Assisted System Identification and Evaluation

3.1. Experiment. The performance of OpenPose and RMPE on MPII and MSCOCO datasets is shown in Figure 8. As can be seen from Figure 8, in the human pose estimation algorithm, the detection speed of OpenPose is 120 times that of RMPE, because OpenPose is composed of the estimated joint points and joint point allocation by the network forward calculation. Among them, the time scale of forward computation is two orders of magnitude larger than the joint point assignment, which dominates. And its time consumption is not affected by the number of people [21]. The experimental results show that OpenPose’s detection speed is 120 times faster than RMPE in a multiperson pose evaluation system. Because OpenPose has network forward computation for estimating joint points and joint point assignments, the forward computation is larger than joint point assignment, and its time consumption is not affected by the number of people. However, the overall detection effect of RMPE for each human joint point and each scale is higher than that of OpenPose. Due to the advantage of more accurate detection of joint points, this article uses the RMPE method for human pose estimation, and the estimated skeleton sequence is input into the action recognition model.

3.2. Comparison of Action Recognition Algorithms. In this experiment, each video ranges from 0.2 seconds to 10 minutes, each video contains a basic action, and the resolution of the video is 1920*1080. And the data are recorded in 15 frames per second, and the movements of the human body joints in the video can be obtained. For the basketball basic action database, the video is first converted into 3D skeleton data, using the above preprocessing method, the 3D information of 18 local joint points is output through human action recognition, and then, the created results are arranged in order [22]. For any initial skeleton data input set to the network, its dimension (18, 3, 300) represents the number of 18 human connection points and the information size of the 3 input connection points. Usually, the space is a three-dimensional coordinate representing the total number of frames of the 300 incoming video, and this sequence is used as the input to the “Human Action Recognition Network.” The 9-layer GCN-TCN module is divided into three parts. The first 3 layers define the number of 64 single-node feature channels, the middle 3 intermediate layers define 128, and the last 3 layers define 256. The ST-GCN recognition algorithm is compared on different datasets, the standard cross-entropy loss function is used as a whole, and the batch size is selected as 16 in the initial training process. Every 5 epoch iterations are reduced to 10% of the original amount,
the initial learning rate is 0.01, a total of 65 epochs are trained, and the recognition rate is verified in the self-built basketball dataset. The model in this article is experimentally compared on the NTU-RGBD dataset, Kinect-skeleton, and the self-built basketball dataset. The experimental results are shown in Figure 9.

As can be seen from Figure 9, since this experiment uses a self-built basketball dataset, the Top1 recognition rate is too low in the initial action recognition results. The parameters need to be adjusted to improve the recognition rate. By adjusting the learning rate lr, this parameter is the gradient of the error to the parameter. If the gradient is positive, it means that it increases, and the loss will also increase. At this time, a positive number can be subtracted to achieve the purpose of reducing the error. If the gradient is negative, it means that it increases, and the loss will decrease accordingly. At this time, a negative number can be subtracted to achieve the purpose of reducing the error. Therefore, after adjusting the parameters for many times, it is finally determined that the lr value is 0.30, the batch is 8, and the epoch is 55, and the recognition accuracy is the highest and the most stable [23]. After modifying the parameters, the value of Top1 of this method is 42.21%, and the value of Top5 is 88.77%. Compared with the Kinect-skeleton dataset, the improvements are 10.61% and 35.09%, respectively. However, compared with the NTU-RGBD dataset, the recognition rate of Top1 is significantly reduced, because the self-built basketball dataset uses manual segmentation of basic actions in the process of collection and processing, and there is a certain error. At the same time, the collected videos have problems such as occlusion, which directly affects the recognition rate.

3.3. Results Display. Through the experimental design in the previous section, two aspects are mainly involved. On the one hand, it is the choice of the preprocessing method for converting the 2D coordinates of the video into 3D coordinates, and on the other hand, it is the use of the action recognition algorithm [24]. By inputting a basketball motion video, the 2D skeleton information is first extracted from the classified video, and then, the 3D skeleton coordinate information is matched in real time through videoPose3D. And by outputting the three-dimensional skeleton video of the player in real time and then processing it through the experimental algorithm, a labeled and classified video is obtained, and the recognition score of each action is given. The basic basketball movements in this article involve shooting, hook layups, dunks, dribbling, running dribbling, and running off the ball. The detailed process and results are shown in Figure 10. In Figure 10(a) represents the preprocessing stage of human action video and 10(b) represents the result of action recognition.

4. Conclusion

In this article, based on the mobile network communication platform and panoramic perspective, the basketball recognition and simulation research is carried out. First, the basketball auxiliary system is constructed and analyzed, which uses the recognition technology of computer vision, and combines the front-end and back-end development technology. To make panorama-based human motion analysis feel beyond isolated actions and poses, contextual information of the environment or objects is integrated. Things such as the context of the environment provide strong knowledge about the type of action, helping to improve recognition accuracy and predict action. With the deepening of action recognition research, recognition technology has made great progress, but it still faces huge challenges. For example, there is still a huge gap between the data images collected by the computer and the biological vision system, and the sensitivity of the biological vision system to the action information is much higher than that of the computer. Based on this, this article studies the action recognition method based on the convolutional neural network. Firstly, the multiperson human pose estimation method is used to extract 2D skeleton information from basketball basic action video data and convert it into 3D skeleton information, and the data are replaced by the input of the convolutional neural network model for action recognition and classification. And using the evaluation
indicators of human pose estimation and action recognition, the experimental results of various action recognition methods and datasets are compared and analyzed, and it is verified that the method has a good recognition effect. There are still some improvements in this paper, including the following: (1) when the athlete’s movements are too fast or too complicated, the recognition results are not ideal. It is necessary to improve the multiperson pose estimation algorithm to extract more accurate human pose estimation information, thereby improving the recognition efficiency. (2) At present, there are no public datasets for professional basketball video data sets. There are 6 types of datasets constructed in this article. There is still a big gap with the basketball video data sets. There are still some improvements in this paper, including the methods and datasets are compared and analyzed, and it is necessary to intensify efforts to collect basic basketball movement videos in the later stage. At the same time, we will discuss action classification with professional coaches for further research on basketball-based action recognition.

Data Availability

The data that support the findings of this study can be obtained from the author upon reasonable request.

Conflicts of Interest

The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] W. Zhu, "Classification accuracy of basketball simulation training system based on sensor fusion and Bayesian algorithm," Journal of Intelligent and Fuzzy Systems, vol. 39, no. 4, pp. 5965–5976, 2020.
[2] A. Castro, M. F. Goethel, E. R. Vieira, P. V. S. Moreira, and M. Gonales, "Effects of wearing an ankle brace on ground reaction force during jumps in basketball game simulation," Revista Brasileira de Medicina do Esporte, vol. 27, no. 2, pp. 218–224, 2021.
[3] X. X. Wang, W. Shuai, Q. Peng et al., "White coat effect in hypertensive patients: the role of hospital environment or physician presence," Journal of the American Society of Hypertension: JASH, vol. 11, no. 8, pp. 498–502, 2017.
[4] W. Pengyu and G. Wanna, "Image detection and basketball training performance simulation based on improved machine learning," Journal of Intelligent and Fuzzy Systems, vol. 40, no. 2, pp. 2493–2504, 2021.
[5] C. Ma, "An improved method for basketball skills and teaching mode optimization based on visual action simulation," Boletin Tecnico/Technical Bulletin, vol. 55, no. 16, pp. 698–704, 2017.
[6] B. Li, Y. Lv, and L. Wang, "An empirical study of basketball technical analysis and tactical training in colleges and universities based on Apriori algorithm," Boletin Tecnico/Technical Bulletin, vol. 55, no. 18, pp. 86–92, 2017.
[7] J. Li and D. Gu, "Research on basketball players’ action recognition based on interactive system and machine learning," Journal of Intelligent and Fuzzy Systems, vol. 40, no. 2, pp. 2029–2039, 2021.
[8] D. Liu, K. Hao, W. Wang et al., "Rv2629 overexpression delays Mycobacterium smegmatis and Mycobacteria tuberculosis entry into log-phase and increases pathogenicity of Mycobacterium smegmatis in mice," Frontiers in Microbiology, vol. 8, no. 1, pp. 2231–2235, 2017.
[9] Z. Hao, X. Wang, and S. Zheng, "Recognition of basketball players’ action detection based on visual image and Harris corner extraction algorithm," Journal of Intelligent and Fuzzy Systems, vol. 40, no. 4, pp. 7589–7599, 2021.
[10] T. Shouyan, "Recognition and simulation of martial arts training behavior based on virtual reality," Agro Food Industry Hi-Tech, vol. 28, no. 1, pp. 2345–2348, 2017.
[11] X. Wu, "Study on the application of human joint force analysis and computer simulation in basketball," Boletin Tecnico/Technical Bulletin, vol. 55, no. 18, pp. 282–289, 2017.
[12] W. Guochen and S. Feihong, "Video analysis method of basketball training assistant based on deep learning theory during COVID-19 spread," Journal of Intelligent and Fuzzy Systems, vol. 39, no. 6, pp. 8747–8755, 2020.
[13] M. S. Ahmed and M. K. Jasim, "The combinatorial effect of healthy educational programs with kinect simulation device on learning some of basketball skills," Indian Journal of Public Health Research & Development, vol. 10, no. 9, pp. 564–565, 2019.
[14] A. Adinoyi, M. Aljamae, and A. Aljlaoud, "The future of broadband connectivity: terrestrial networks vs. satellite constellations," International Journal of Communications, Network and System Sciences, vol. 15, no. 05, pp. 53–66, 2022.
[15] Z. Liu, R. Wang, and D. Tang, "Extending labeled mobile network traffic data by three levels traffic identification fusion," Future Generation Computer Systems, vol. 88, pp. 453–466, 2018.
[16] E. Azhikannickal, "Sports, smartphones, and simulation as an engaging method to teach projectile motion incorporating air resistance," The Physics Teacher, vol. 57, no. 5, pp. 308–311, 2019.
[17] B. J. Han and J. Y. Sim, "Saliency detection for panoramic landscape images of outdoor scenes," Journal of Visual Communication and Image Representation, vol. 49, pp. 27–37, 2017.
[18] J. Kim, "Panoramic image communication for mobile application using content-aware image resizing method," International Journal of Advanced Science, Engineering and Information Technology, vol. 7, no. 2, pp. 338–339, 2017.
[19] M. Mitra, "Satellite communication," Iete Journal of Education, vol. 38, no. 1, pp. 3–10, 2018.
[20] S. Kavitha and S. Sutha, "Pollution monitoring and controlling system using global system for mobile communication network," Bonfring International Journal of Software Engineering and Soft Computing, vol. 9, no. 2, pp. 40–42, 2019.
[21] Perceptive Mobile Networks, "Cellular networks with radio vision via joint communication and radar sensing," IEEE Vehicular Technology Magazine, vol. 16, no. 2, pp. 20–30, 2021.
[22] L. Liu, Z. Yan, and Q. Liu, "Panoramic visual tracking based on adaptive mechanism," Journal of Visual Communication and Image Representation, vol. 57, pp. 99–106, 2018.
[23] Y. Zhang and F. Huang, "Panoramic visual SLAM technology for spherical images," Sensors, vol. 21, no. 3, pp. 705–706, 2021.
[24] I. A. Kudinov, M. B. Nikiforov, and I. S. Kholopov, "Strategies for generating panoramic video images without information about scene correspondences for multispectral distributed aperture systems," Computer Optics, vol. 45, no. 4, pp. 589–599, 2021.