

Research Article

Sensor Fusion Basketball Shooting Posture Recognition System Based on CNN

Jingjin Fan,¹ Shuoben Bi ,^{1,2} Guojie Wang,² Li Zhang,¹ and Shilei Sun²

¹Research Institute of History for Science and Technology, Nanjing University of Information Science & Technology, Nanjing 210044, China

²School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China

Correspondence should be addressed to Shuoben Bi; bishuoben@163.com

Received 20 December 2020; Revised 7 March 2021; Accepted 12 March 2021; Published 30 March 2021

Academic Editor: Bin Gao

Copyright © 2021 Jingjin Fan 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.

In recent years, with the development of wearable sensor devices, research on sports monitoring using inertial measurement units has received increasing attention; however, a specific system for identifying various basketball shooting postures does not exist thus far. In this study, we designed a sensor fusion basketball shooting posture recognition system based on convolutional neural networks. The system, using the sensor fusion framework, collected the basketball shooting posture data of the players' main force hand and main force foot for sensor fusion and used a deep learning model based on convolutional neural networks for recognition. We collected 12,177 sensor fusion basketball shooting posture data entries of 13 Chinese adult male subjects aged 18–40 years and with at least 2 years of basketball experience without professional training. We then trained and tested the shooting posture data using the classic visual geometry group network 16 deep learning model. The intratest achieved a 98.6% average recall rate, 98.6% average precision rate, and 98.6% accuracy rate. The intertest achieved an average recall rate of 89.8%, an average precision rate of 91.1%, and an accuracy rate of 89.9%.

1. Introduction

Basketball is one of the most popular sports with a large fan base worldwide. As a competitive sport, basketball requires two teams of players to use various technical actions to compete with each other. A basketball game includes various technical statistics, such as score, rebound, assist, block, and steal, among which score is the central aspect that decides which team is the winner and loser [1–3]. The score in a basketball match is accomplished through players shooting the ball into the basket, and, thus, shooting plays an important tactical role in this game.

In recent years, with the rapid development of wearable sensor technology and the increased demand for basketball worldwide, many researchers have used wearable devices integrated with an internal measurement unit (IMU) to study basketball shooting [4–6]. Bai et al. [7] used Microsoft Band and weSport systems to collect data from two basketball players on both attack and defense, and they used a support vector machine (SVM) to effectively distinguish shooting

and defense in basketball games. Aacikmese et al. [8], using an IMU placed on the arm, classified the six technical movements (forward-backward dribbling, left-right dribbling, regular dribbling, two-handed dribbling, shooting, and lay-up) in basketball by SVM. Zhao et al. [9] used four IMUs placed on the left and right upper arms and forearms to collect basketball technical movement data, using SVM to identify dribbling, passing, catching, and shooting. These studies, which use sensors on the arms, address basic shooting postures but ignore composite shooting postures. Composite shooting postures are shooting postures that consist of a series of hand and foot movements [10]. It is not sufficient to study composite shooting postures using only arm sensor data.

Shooting is a technical movement that requires physical coordination. In the shooting process, the movement of the feet is as important as the movement of the arms [1, 10]. Shi et al. [11] used smart insoles integrated with IMUs to distinguish between dribbling, jumping, and turning around during basketball. Peng et al. [12] also used smart insoles to

study the sideslip, back, cross, jab, and jump steps in basketball. These studies, which concern footstep movement in basketball using smart insoles integrated with IMUs, provide a basis for us to carry out research on composite shooting postures.

After more than 100 years of development, basketball has developed many complex and delicate technical movements, such as shooting. Recognizing these technical movements requires powerful recognition tools. SVM is a simple and robust algorithm and is widely used in basketball technical movement recognition [7–9]. However, SVM requires feature extraction and cannot be applied to large-scale training samples. Convolutional neural networks (CNNs) have solved these problems. A convolutional neural network is one of the representative algorithms of deep learning. Research on CNNs started in the 1980s and 1990s. In the twenty-first century, with the introduction of deep learning and the development of computer hardware capable of supporting deep learning, there has been a rapid development of CNNs, and these CNNs have worked well in the fields of computer vision and natural language processing [13–15]. With the development of CNNs, many researchers have studied sports using multiple IMUs combined with a CNN model of deep learning. Lee et al. [16] used the CNN long short-term memory model of deep learning to classify six squat positions (one correct and five incorrect); Kautz et al. [17] used a deep CNN to classify 10 types of beach volleyball technical movements. The effectiveness of these studies illustrates the great potential of deep learning models based on CNNs in the field of sports technical movement recognition.

Although there are many sports monitoring systems based on IMUs [6, 18, 19], there is still no system based on deep learning models to recognize a variety of basketball shooting postures. This study proposes a sensor fusion basketball shooting posture recognition system based on a CNN to recognize multiple types of basketball shooting postures. The main features of this study are as follows:

- (1) A sensor fusion framework dedicated to basketball shooting postures is designed to collect shooting posture data and perform sensor data fusion
- (2) 10 types of sensor fusion basketball shooting posture datasets are established, which can be used in related studies
- (3) The 10 types of sensor fusion basketball shooting posture datasets were trained and tested using the classic visual geometry group network 16 (VGG16) deep learning model based on CNN, and the results verified the effectiveness of the deep learning model in shooting posture recognition

The remainder of this paper is organized as follows. Section 2 briefly summarizes the system framework and methods of data collection, fusion, and classification. Section 3 presents the experiments and results. Section 4 discusses the results of this study. Finally, conclusions are presented in Section 5.

2. Materials and Methods

2.1. System Hardware and Software Design. The sensor fusion basketball shooting posture recognition system consists of two independent wireless sensor modules, a USB dongle, and a laptop computer, as shown in Figure 1. The wireless sensor module is composed of an IMU (mpu-9250, including accelerometer and gyroscope) and a microcontroller unit (MCU, Nordic nrf52832, including Bluetooth functionality). The module is powered by an external 500 mA 3.3 V lithium battery. The IMU is responsible for collecting the players' raw shooting posture data, including accelerometer data and gyroscope data. The sampling rate was 100 Hz, and the collected data were transmitted to the MCU through I2C. The MCU is the core component of the wireless sensor module and is responsible for transmitting the raw shooting posture data from the IMU to the USB dongle via Bluetooth. The USB dongle includes Bluetooth and USB human interface device (HID) functions; it is responsible for transmitting the raw shooting posture data received from the wireless sensor modules to a laptop through a USB HID. The laptop contains a data-processing software developed by MATLAB, which is responsible for receiving, displaying, and fusing the raw shooting posture data to form the sensor fusion basketball shooting posture datasets.

2.2. Sensor Fusion Framework. Consider right-handed players as an example. When a basketball player shoots, his right hand is the main force hand, and at the same time, his left foot is the main force foot. The main force hand and main force foot perform the main tasks in basketball shooting. They make the shooting postures stable and can best reflect the characteristics of the shooting postures [10]. Therefore, the posture data of the right hand and left foot of right-handed players are key data. Correspondingly, the key data of the left-handed players were generated from the left hand and right foot. Modern basketball has several types of shooting posture. For basic shooting postures, the main force hand sensor data reflects the characteristics of the shooting postures. For composite shooting postures, such as stop jump shots and gather step shots, the main force hand sensor data cannot fully reflect the characteristics of the postures. However, the main force hand sensor data fused with the main force foot sensor data can fully reflect the characteristics of composite shooting postures. Therefore, this study proposes a sensor fusion framework for basketball shooting postures. This framework collects accelerometer and gyroscope data using wireless sensor modules placed on the main force hand and the main force foot of the player. Then, the data are fused to form the input of the deep learning model for shooting posture classification. The sensor fusion framework proposed in this study can classify a variety of complex shooting postures without increasing the number of sensors.

As shown in Figure 2, the sensor fusion framework proposed in this study consists of three steps: (1) shooting posture data collection, (2) data alignment and merge, and (3) data segmentation and exclusion. First, we synchronized the two independent wireless sensor modules and placed them on the player's main force hand and main force foot.

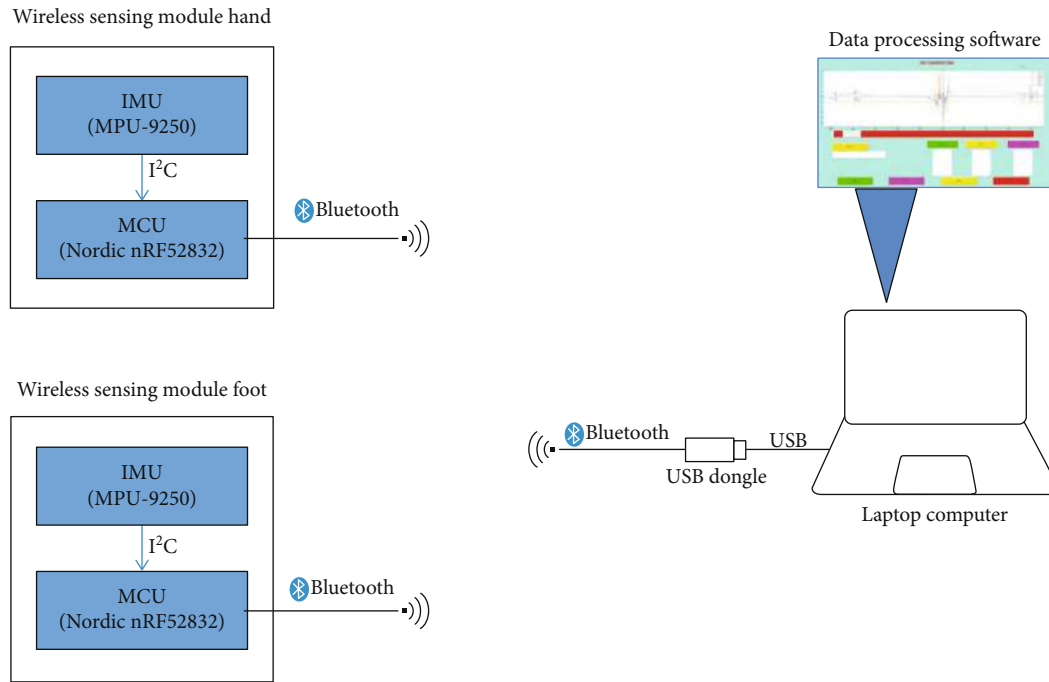


FIGURE 1: Sensor fusion basketball shooting posture recognition system.

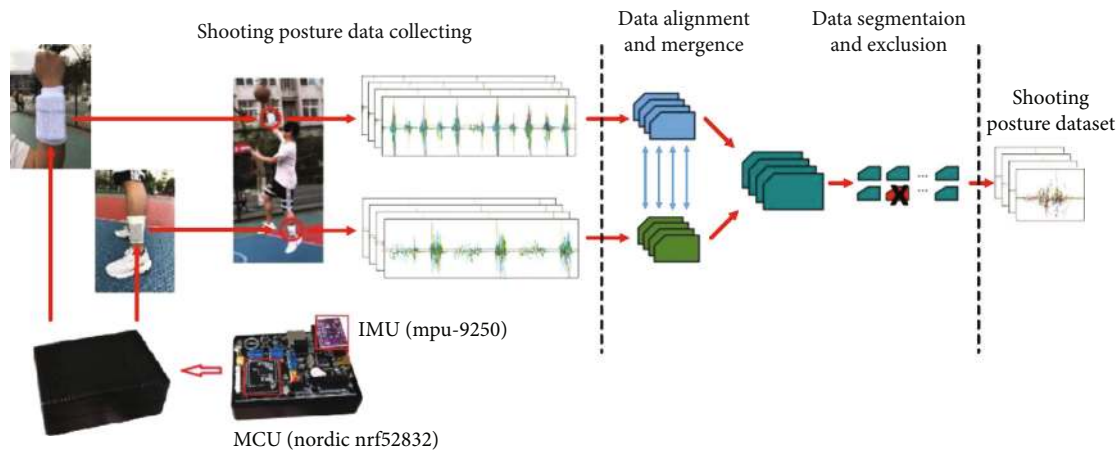


FIGURE 2: Sensor fusion framework.

Then, we collected the shooting posture data, which contain timestamps and transmit them to the laptop computer, stored in two separate files. Because our sampling rate is 100 Hz, the timestamp is in units of 10 ms. Second, the data in the two data files are aligned according to the timestamps and merged into a sensor fusion data file, as shown in Algorithm 1. File *_H*, file *_F*, and file *_M* represent the main force hand sensor data file, main force foot sensor data file, and sensor fusion data file, respectively. Owing to data loss in the wireless sensor module and other reasons, the shooting posture data of the main force hand and main force foot did not match, and hence, the sensor fusion data frequency suffered a loss of 1.17%. However, the frequency reduction did not affect the recognition of shooting postures. Finally, we divided the sensor fusion data file into independent

shooting posture data entries, removed the erroneous posture data, and stored them in the sensor fusion basketball shooting posture dataset, as shown in Algorithm 2, where matrix(*i*) represents the *i*th shooting posture data matrix. We marked the data generated due to sensor misplacement or incorrect shooting posture in the experimental stage and deleted it in this stage. Thereafter, the sensor fusion basketball shooting posture dataset was finally formed.

2.3. Classification Model. The VGG16 [20–22] deep learning model based on CNN [23–25] is a model for image recognition proposed by the Visual Geometry Group of the University of Oxford in 2014. This model participated in the 2014 ImageNet Image Classification and Positioning Challenge and achieved excellent results. Because the

```

1. Open file_H, file_F, and file_M
2. Send file_H first record to record_H, send file_F first record to record_F
3. while (file_H NOT end) AND (file_F NOT end) do
4. if record_H.timestamp == record_F.timestamp do
5. Merge record_H and record_F to file_M
6. Send file_H next record to record_H, send file_F next record to record_F
7. else if record_H.timestamp > record_F.timestamp do
8. Delete record_F, send file_F next record to record_F
9. else if record_H.timestamp < record_F.timestamp do
10. Delete record_H, send file_H next record to record_H
11. end if
12. end while
13. Close file_H, file_F, and file_M

```

ALGORITHM 1: Data alignment and mergence.

VGG16 deep learning model showed excellent performance in image classification, and the VGG16 model with one-dimensional convolution kernels had been used to classify the one-dimensional data obtained by using the accelerometer and gyroscope [26], in this study, we used the one-dimensional convolution kernels VGG16 deep learning model to classify sensor fusion basketball shooting postures.

The structure of the VGG16 deep learning model mainly includes convolutional layer, max pooling layer, and fully connected layer, as shown in Figure 3. The function of the convolutional layer, which consists of several convolutional units, is to extract different features of the input data. Adding a greater number of convolutional layers means a greater number of complex features can be extracted. The working mode of the convolutional layer can be expressed by Equations (1), (2), and (3):

$$D = f(S * C), \quad (1)$$

$$S = [S_1, S_2, \dots, S_m]^T = \begin{bmatrix} s_{11} \cdots s_{1n} \\ \vdots \quad \ddots \quad \vdots \\ s_{m1} \cdots s_{mn} \end{bmatrix}_{m \times n}, \quad (2)$$

$$C = [C_1^T, C_2^T, \dots, C_k^T]_{l \times k} (l < n), \quad (3)$$

where S represents the vector of sample data, C represents the vector of the convolution kernel, m is the number of sample data, n is the number of input features, l is the length of the convolution kernel, and k is the number of convolution kernels. The result of the convolution operation of the i th convolution kernel is shown in Equation (4).

$$D_i = \begin{bmatrix} S_1 * C_i^T \\ \vdots \\ S_m * C_i^T \end{bmatrix}_{m \times n}, \quad (4)$$

where $i = 1, 2, 3, \dots, k$. Because we use padding, the width of the vector after the convolution operation is n .

```

1. Open file_M
2. Send file_M first record to record_M
3. while file_M NOT end do
4. if record_M is shooting interval do
5. Delete record_M
6. else if record_M is shooting posture do
7. Send record_M to the shooting matrix(i)
8. if shooting posture is end do
9. if matrix(i) is not error do
10. Store matrix(i) to posture dataset
11. endif
12. Clear matrix(i)
13. end if
14. end if
15. Send file_M next record to record_M
16. end while
17. Close file_M

```

ALGORITHM 2: Data segmentation and exclusion.

The max pooling layer is mainly used for reducing feature dimensionality, compressing the number of data and parameters, reducing overfitting, and improving the fault tolerance of the model. The working mode of the max pooling layer can be expressed by Equations (5) and (6), where a is the stride.

$$Q = \max \text{pool}(D) = [Q_1, Q_2, \dots, Q_k], \quad (5)$$

$$Q_j = \begin{bmatrix} Q_{11} & \cdots & Q_{1(n-a+1)} \\ \vdots & \ddots & \vdots \\ Q_{m1} & \cdots & Q_{m(n-a+1)} \end{bmatrix}_{m \times (n-a+1)}. \quad (6)$$

The fully connected layer mainly plays the role of classification, which is used to integrate and map the distributed feature representation extracted by the convolutional layer and the max pooling layer to the sample label space. The output of the fully connected layer is the final classification result.

The weight initialization of the convolutional layer and fully connected layer uses the Kaiming method [27], which can accelerate the convergence speed of the model.

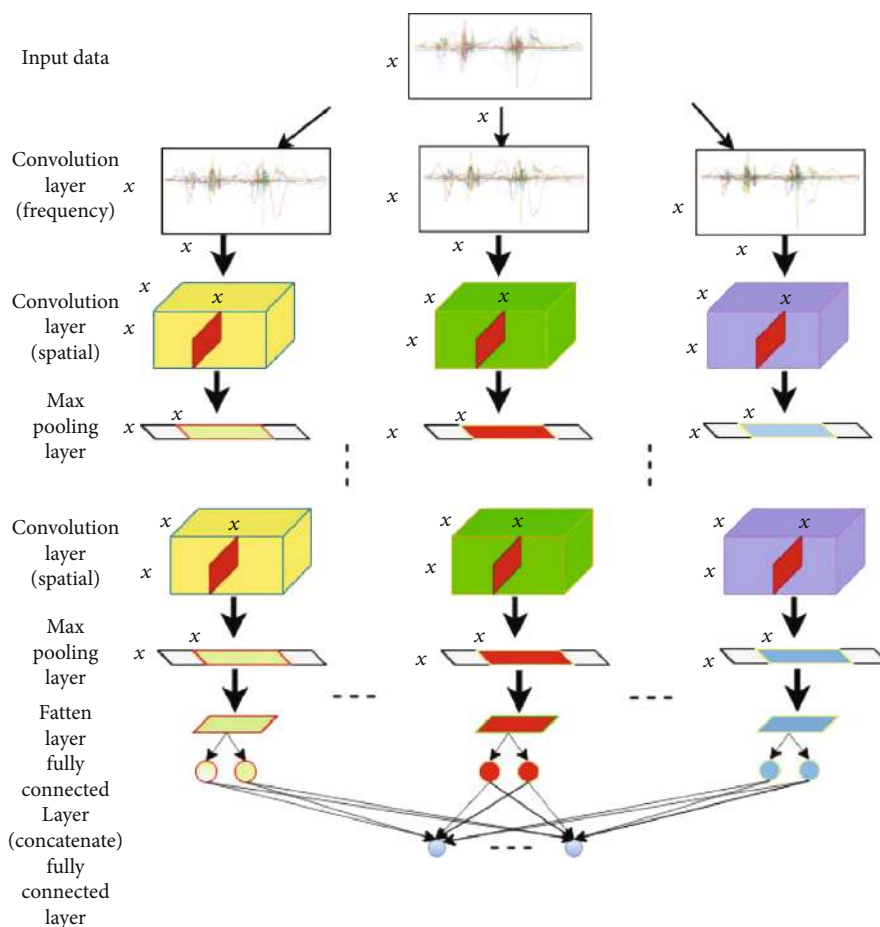


FIGURE 3: Structure of the visual geometry group network 16 (VGG16) deep learning model.

TABLE 1: Parameter settings of the visual geometry group network 16 (VGG16) deep learning model.

Parameters	Value
CNN layer weight initialization	Kaiming
Full connection layer weight initialization	Kaiming
Data standardization	Z-Score
Optimizer	Adam
Initial learning rate	1E-3
Loss function	Cross entropy
Network layer activation function	Rectified linear unit (ReLU)
Pooling method	Maximum pooling
Training rounds	20
Dropout rate	0.5
Batch size	200
Padding	[1]

The Z-Score method [28], which can convert datasets of different measurements into a unified measurement of Z-Score for comparison, is adopted for data standardization. The model uses the Adam optimizer [29], which has the

TABLE 2: Characteristics of the subjects.

Parameters	Values
Age (years)	28.5 ± 9.5
Height (cm)	179 ± 14
Weight (kg)	80.5 ± 21.5
Experience (years)	12 ± 9

advantages of simple implementation, high calculation efficiency, and lower memory requirements, and it is suitable for large-scale data and parameter scenarios. It is often used as the optimization algorithm for stochastic gradient descent (SGD). The minibatch gradient descent algorithm adopted for the model has the high speed of the SGD algorithm as well as the stability of the batch algorithm, which is suitable for deep learning models that need to process large amounts of data [25]; in the proposed model, the batch size is set to 200. The specific parameter settings are listed in Table 1.

3. Experimental Method and Results

3.1. Experimental Method. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the IEC for Clinical Research of

TABLE 3: Basketball shooting postures.

Shooting postures	Explanation
Gather step shot	Outside the painted area, dribble one or two times, and then step back and shoot with both hands.
Hook shot	Inside the painted area, from the side of the basket, do not take off, hook shoot with one hand.
Free throw	After the free throw line, do not take off, shoot with both hands.
Stop jump shot	Outside the painted area, dribble, stop abruptly and jump, and then shoot with both hands.
Pump fake	Inside the painted area, make a fake shot with both hands, and then shoot with both hands.
Inside shot	Inside the painted area, under the basket, facing the basket, do not take off, shoot with both hands.
Jettison throw	From the three-point line, dribble forward, jump, and throw the ball into the basket with one hand.
Lay-up	From the three-point line, dribble forward, and perform a one-hand lay-up.
Jump shot	Outside the painted area, without dribbling, jump and shoot with both hands.
Spin jumper	Inside three-point line, dribble one or two times, and then turn around and shoot with both hands.

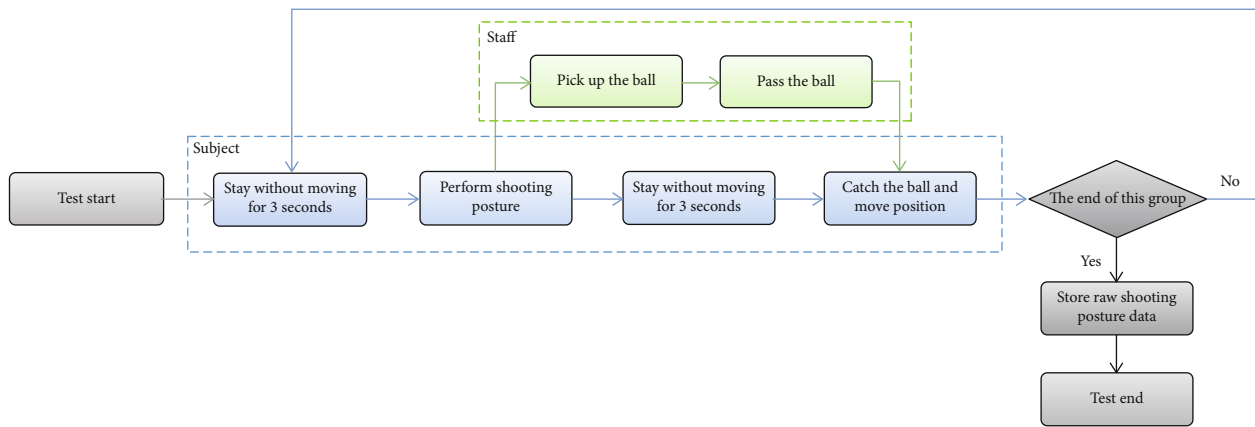


FIGURE 4: Shooting posture test flowchart.

Zhongda Hospital, affiliated with Southeast University (Project identification code: 2020ZDSYLL151-P01).

A total of 13 Chinese male adults (age 28.5 ± 9.5 years, height 179 ± 14 cm, weight 80.5 ± 21.5 kg) were selected as subjects. Although they had basketball experience (12 ± 9 years), they were not professional players and had no professional training, as reported in Table 2. All subjects were right-handed players. There were two centers (C), two power forwards (PF), three small forwards (SF), three shooting guards (SG), and three point guards (PG). Among them, five subjects had been trained as part of a college team. All subjects gave their informed consent for inclusion before they participated in the study. The subjects were verbally informed of the experiment process and precautions to be taken before the start of the experiment.

We chose 10 types of basketball shooting postures [30–33] for the experiment, as summarized in Table 3. These postures included five basic shooting postures: hook shot, free throw, inside shot, lay-up, and jump shot. In addition, we chose five types of composite shooting postures: gather step shots, stop jump shots, pump fakes, jettison throws, and spin jumpers. The stop jump shot, pump fake, and spin jumper are frequently used in basketball. The gather step shots and jettison throws are new introductions to the sports that have become increasingly popular in recent years.

The experiment was conducted in the basketball court of the Nanjing University of Information Science and Technology. The subjects repeated each of the 10 types of basketball shooting postures, as shown in Tables 3, 50–150 times. Each shooting posture was divided into 1–4 groups according to the physical strength of the subjects, with 25–150 shooting posture cycles in each group. At the beginning of each shooting posture cycle, the subjects held the ball without moving for 3 s and then performed the corresponding shooting posture. When the shooting posture was finished, they did not move until after 3 s had passed. Immediately after the shooting posture was completed, the staff picked up the ball and passed it to the subject after the subject moved. If the testing of the group was not complete, the next shooting posture cycle was started after the subject received the ball. If the testing of the group was complete, the data-processing software stored the raw shooting posture data, which contained 25–150 shooting posture cycles for sensor data fusion. The shooting posture test process is presented in Figure 4.

Finally, 10 types of sensor fusion basketball shooting posture datasets of 13 subjects, a total of 12,210 shooting posture data entries, including 12,177 valid data entries, were formed. The datasets included 1,210 gather step shots, 1,228 hook shots, 1,209 free throws, 1,223 stop jump shots, 1,221 pump fakes, 1,225 inside shots, 1,218 jettison throws, 1,216 lay-

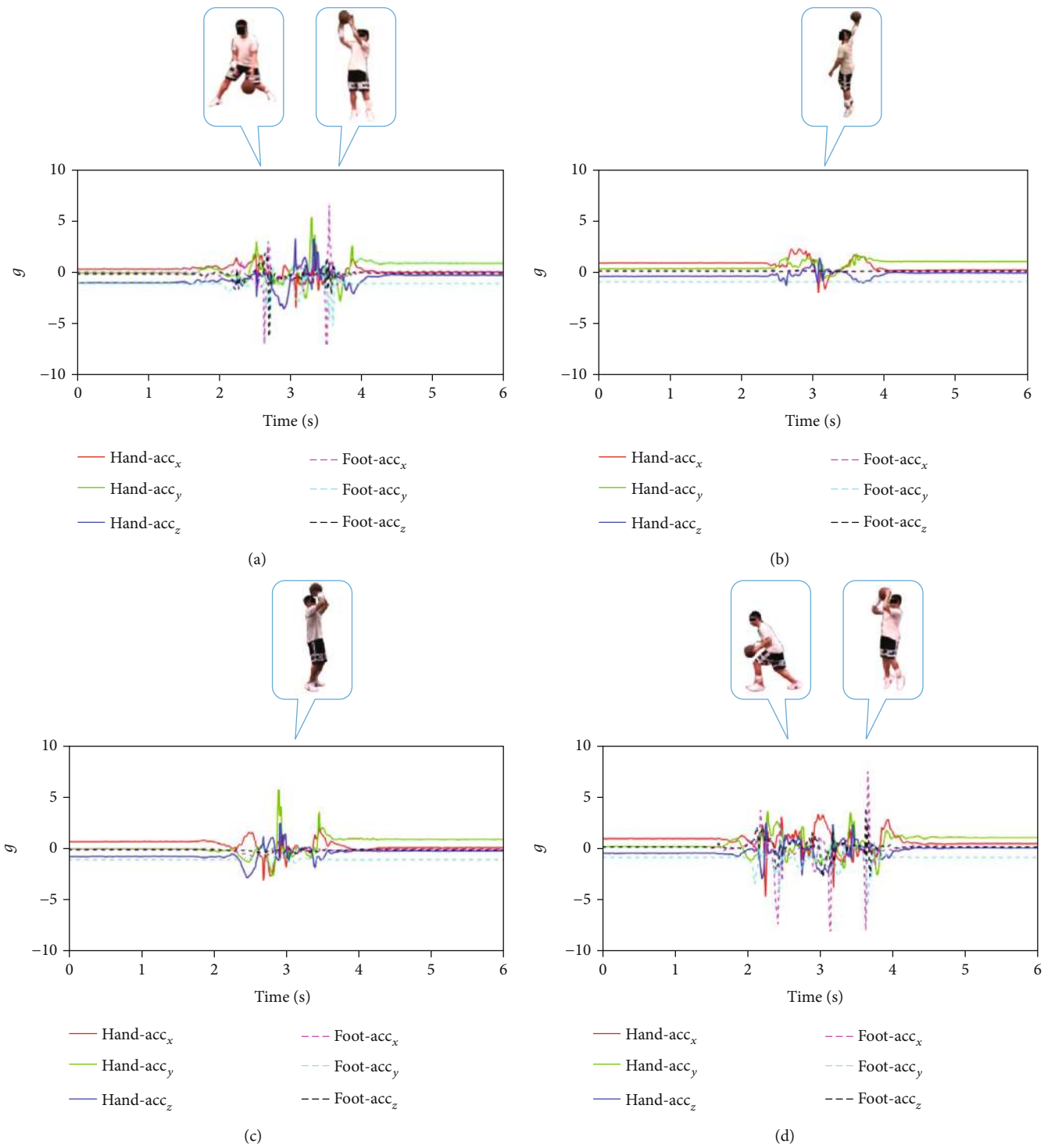


FIGURE 5: Continued.

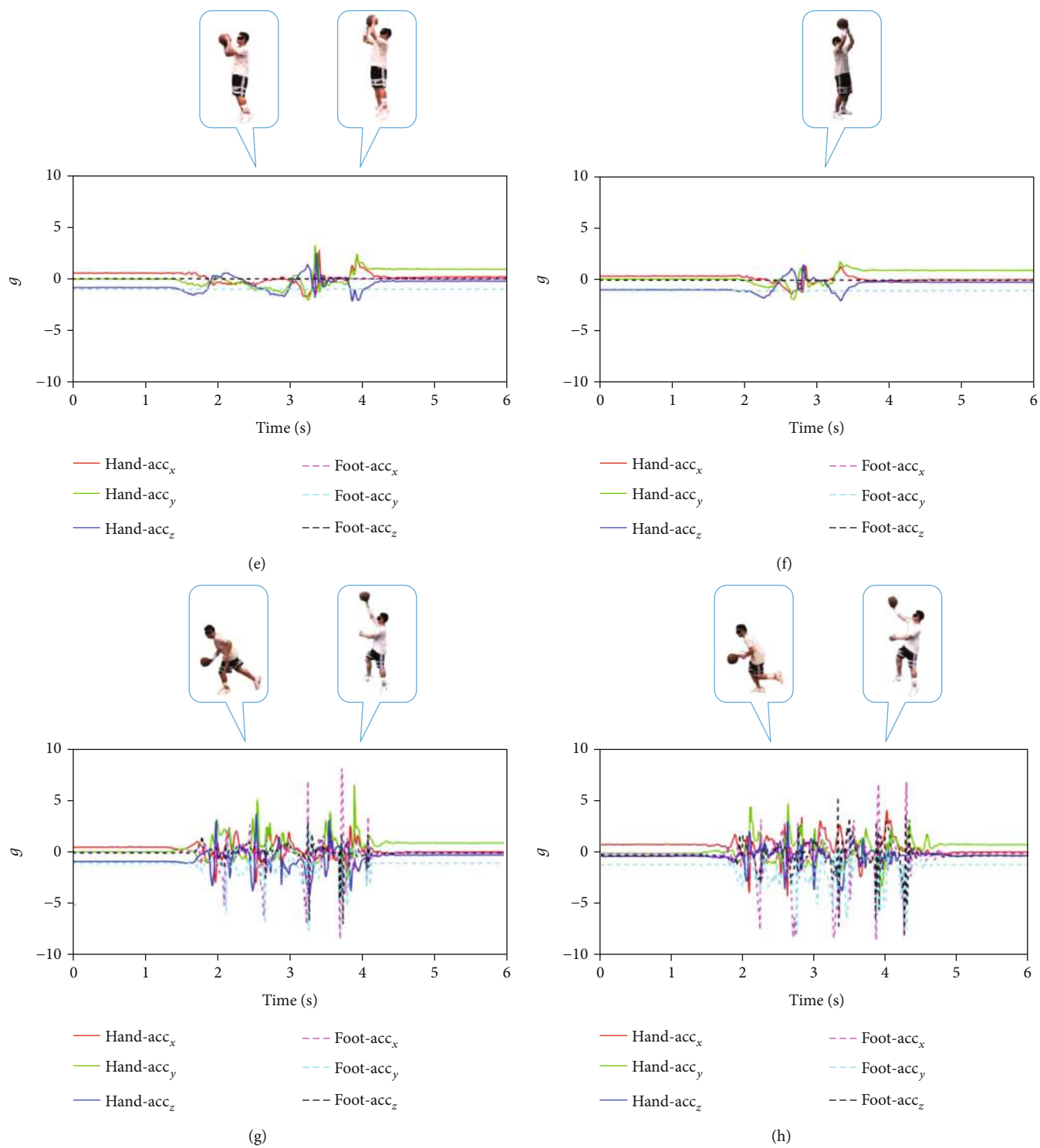


FIGURE 5: Continued.

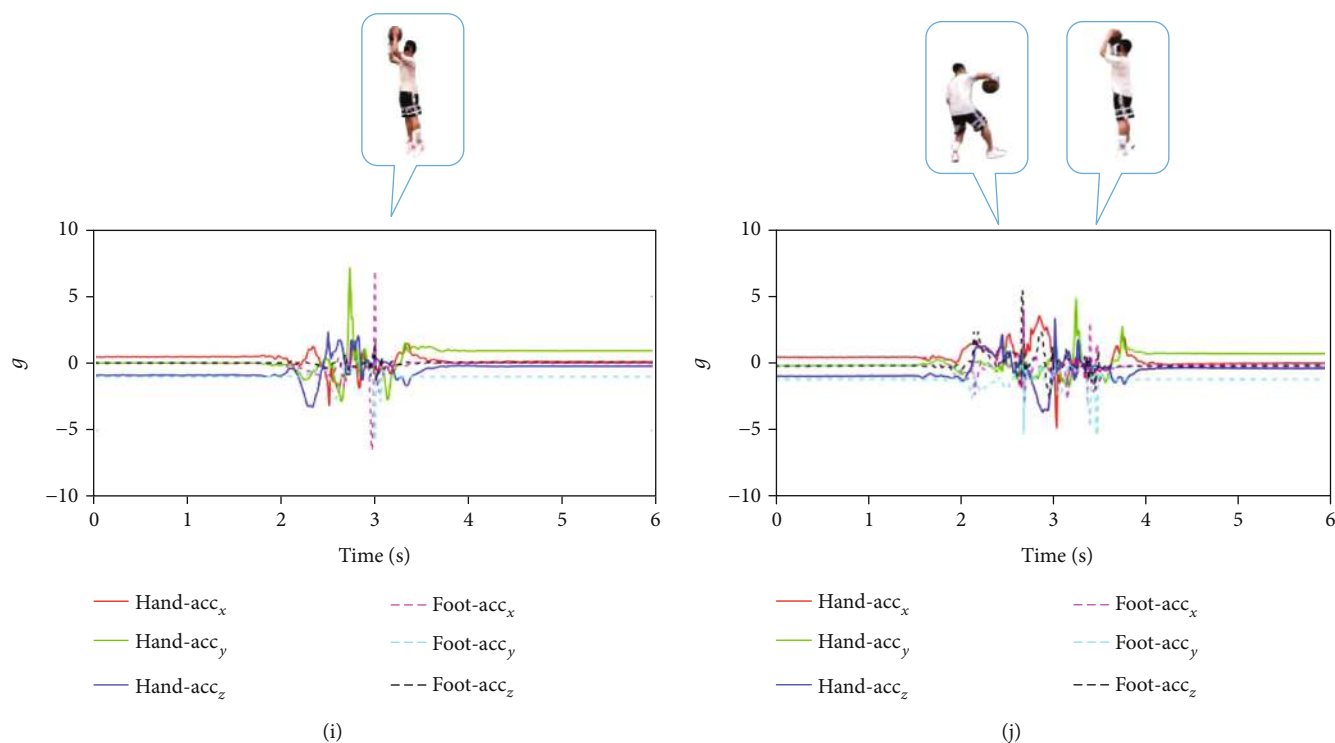


FIGURE 5: Accelerometer data waveforms of the 10 types of sensor fusion basketball shooting postures: (a) gather step shot, (b) hook shot, (c) free throw, (d) stop jump shot, (e) pump fake, (f) inside shot, (g) jettison throw, (h) lay-up, (i) jump shot, and (j) spin jumper.

ups, 1,207 jump shots, and 1,220 spin jumpers. The accelerometer data waveforms of the 10 types of sensor fusion basketball shooting postures are shown in Figure 5.

3.2. Classification. In this study, intra- and intertraining and testing methods were used for the sensor fusion basketball shooting posture datasets. Both methods were carried out on a computer configured with a Core i5-9400 CPU, 32 GB memory, and a GeForce GT730 graphics card. The operating system was Windows 10 Home, and the model was implemented using the MATLAB 2019b Deep Learning Toolbox.

3.2.1. Intratraining and Testing. All 12,177 sensor fusion basketball shooting posture data entries were randomly arranged, and the training and test datasets were designed with an 8:2 ratio, including 9,741 data entries in the training dataset and 2,436 data entries in the test dataset. The training dataset was used to train the model, and the test dataset was used to test the model. Figure 6 presents a comparison between the loss rate and accuracy rate of the intratraining process. As the loss rate decreases, the accuracy rate continuously increases, demonstrating the continuous improvement of the training model. Figure 7 presents the confusion matrix of 10 types of sensor fusion basketball shooting posture test dataset classified by the intratest. The row variables of the matrix represent the recall rate and false negative rate, and the column variables represent the precision rate and false discovery rate.

3.2.2. Intertraining and Testing. The sensor fusion basketball shooting posture data of 13 subjects were randomly arranged;

the data of 11 subjects were used to form the training dataset, and the data of two subjects were used to form the test dataset. The total number of training data entries was 10,126, and the total number of test data entries was 2,051. The training dataset was used to train the model, and the test dataset was used to test the model. Figure 8 shows a comparison between the loss rate and the accuracy rate of the intertraining process. Figure 9 presents the confusion matrix of the 10 types of sensor fusion basketball shooting postures test dataset classified by the intertest.

3.3. Results and Analysis. Figure 10 depicts the t-SNE diagram of the intratest dataset. The t-SNE diagram shows the distribution characteristics of the data intuitively by reducing high-dimensional data to two-dimensional data. From the t-SNE diagram of the intratest dataset, it can be observed that the lay-up, jettison throw, and stop jump shot are easily confused, as well as hook shot, inside shot, and free throw.

Table 4 summarizes the recall rate, precision rate, average recall rate, and average precision rate of the intratest classification results. The classification results of the intratest reveal that the average recall rate was 98.6%, and the maximum recall rate was 100% for the jump shot and spin jumper. The minimum recall rate was 96% for the gather step shot. The average precision rate was 98.6%; the maximum precision rate was 100% for the hook shot, and the minimum precision rate was 97.2% for the jump shot. The above data indicate that although the t-SNE diagram shows that there are easily confused shooting postures in the intratest dataset, the sensor fusion basketball shooting posture recognition

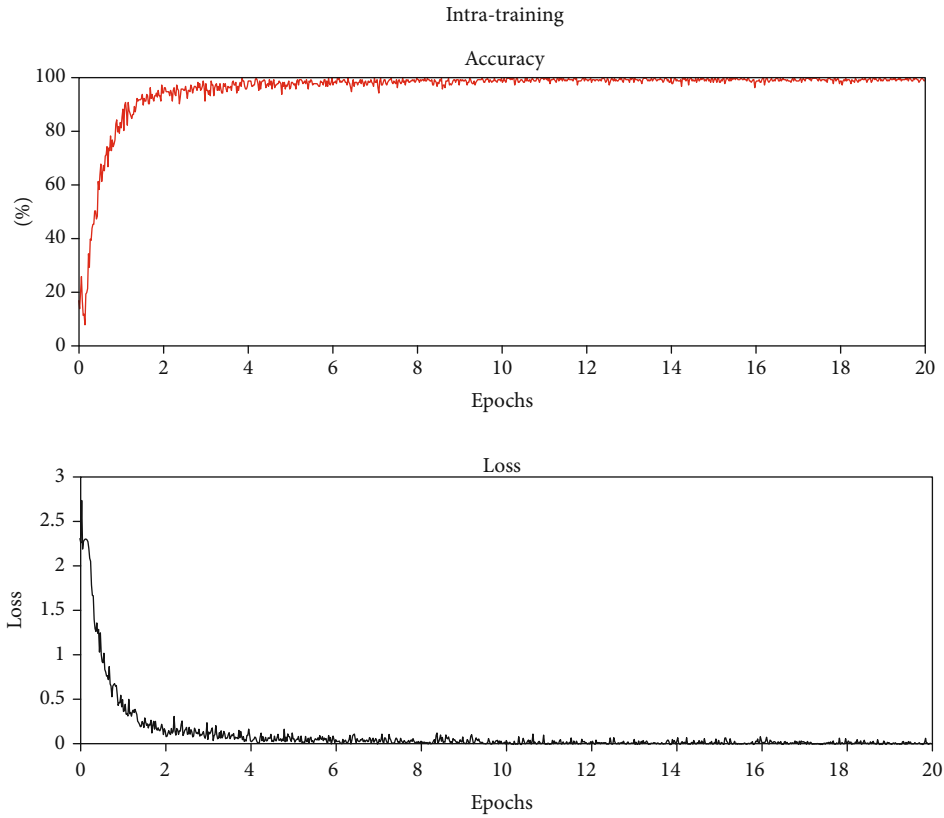


FIGURE 6: Accuracy rate curve and loss function value curve of the intratraining process.

		Intra-test											
Output class	Gather step shot	239 9.8%	0 0.0%	0 0.0%	1 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.2% 0.8%	
	Hook shot	0 0.0%	237 9.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	
	Free throw	0 0.0%	0 0.0%	209 8.6%	0 0.0%	3 0.1%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.1% 1.9%	
	Stop jump shot	4 0.2%	0 0.0%	0 0.0%	253 10.4%	0 0.0%	0 0.0%	1 0.0%	1 0.0%	0 0.0%	0 0.0%	97.7% 2.3%	
	Pump fake	0 0.0%	0 0.0%	1 0.0%	0 0.0%	221 9.1%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.1% 0.9%	
	Inside shot	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	282 11.6%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%	
	Jettison throw	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	240 9.9%	3 0.1%	0 0.0%	0 0.0%	98.4% 1.6%	
	Lay-up	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	1 0.0%	242 9.9%	0 0.0%	0 0.0%	99.2% 0.8%	
	Jump shot	3 0.1%	0 0.0%	3 0.1%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	243 10.0%	0 0.0%	97.2% 2.8%	
	Spin jumper	3 0.1%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	236 9.7%	97.9% 2.1%	
			96.0% 4.0%	99.2% 0.8%	98.1% 1.9%	98.4% 1.6%	98.2% 1.8%	99.3% 0.7%	98.8% 1.2%	98.0% 2.0%	100% 0.0%	100% 0.0%	98.6% 1.4%
			Gather step shot	Hook shot	Free throw	Stop jump shot	Pump fake	Inside shot	Jettison throw	Lay-up	Jump shot	Spin jumper	
			Target class										

FIGURE 7: Classification confusion matrix of the intratest.

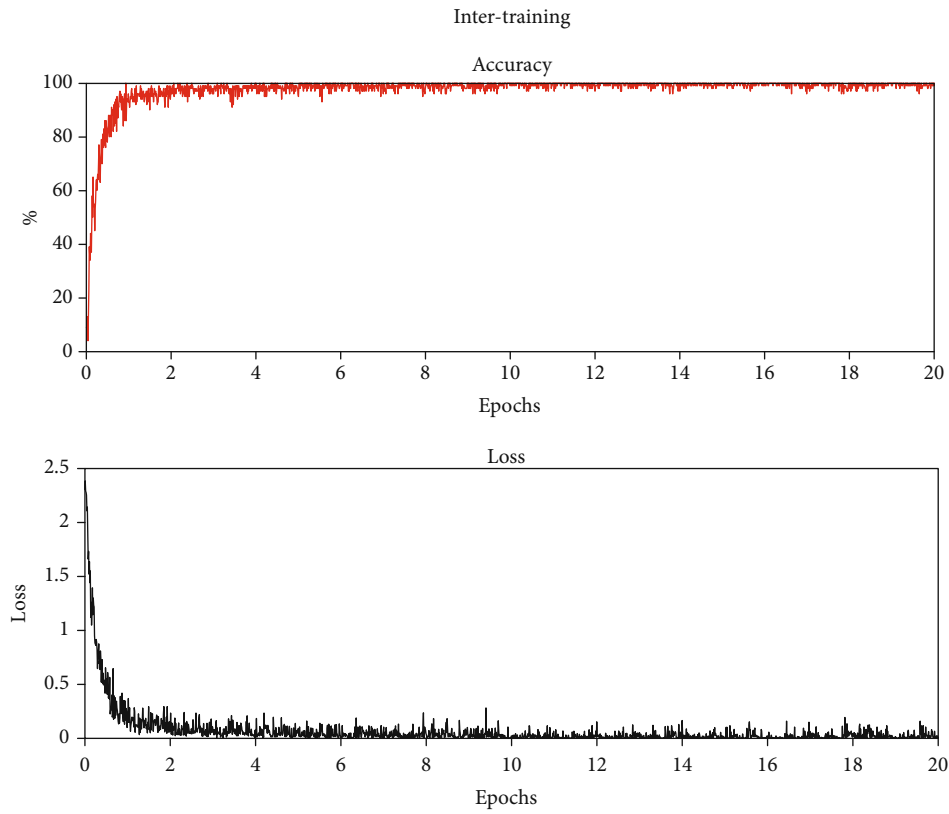


FIGURE 8: Accuracy rate curve and loss function value curve of the intertraining process.

		Inter-test										
Output class	Gather step shot	202 9.8%	0 0.0%	0 0.0%	6 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	97.1% 2.9%
	Hook shot	0 0.0%	209 10.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	Free throw	0 0.0%	1 0.0%	210 10.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	36 1.8%	0 0.0%	85.0% 15.0%
	Stop jump shot	0 0.0%	0 0.0%	0 0.0%	219 10.7%	0 0.0%	0 0.0%	35 1.7%	0 0.0%	0 0.0%	0 0.0%	86.2% 13.8%
	Pump fake	0 0.0%	3 0.1%	4 0.2%	0 0.0%	182 8.9%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	95.8% 4.2%
	Inside shot	0 0.0%	74 3.6%	20 1.0%	0 0.0%	0 0.0%	182 8.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	65.9% 34.1%
	Jettison throw	0 0.0%	0 0.0%	0 0.0%	7 0.3%	0 0.0%	0 0.0%	130 6.3%	1 0.0%	0 0.0%	0 0.0%	94.2% 5.8%
	Lay-up	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	18 0.9%	132 6.4%	0 0.0%	0 0.0%	87.4% 12.6%
	Jump shot	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	94 4.6%	0 0.0%	98.9% 1.1%
	Spin jumper	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	283 13.8%	100% 0.0%
			100 0.0%	72.8% 27.2%	89.4% 10.6%	94.0% 6.0%	100% 0.0%	100% 0.0%	71.0% 29.0%	99.2% 0.8%	71.8% 28.2%	100% 0.0%
	Target class	Gather step shot	Hook shot	Free throw	Stop jump shot	Pump fake	Inside shot	Jettison throw	Lay-up	Jump shot	Spin jumper	

FIGURE 9: Classification confusion matrix of the intertest.

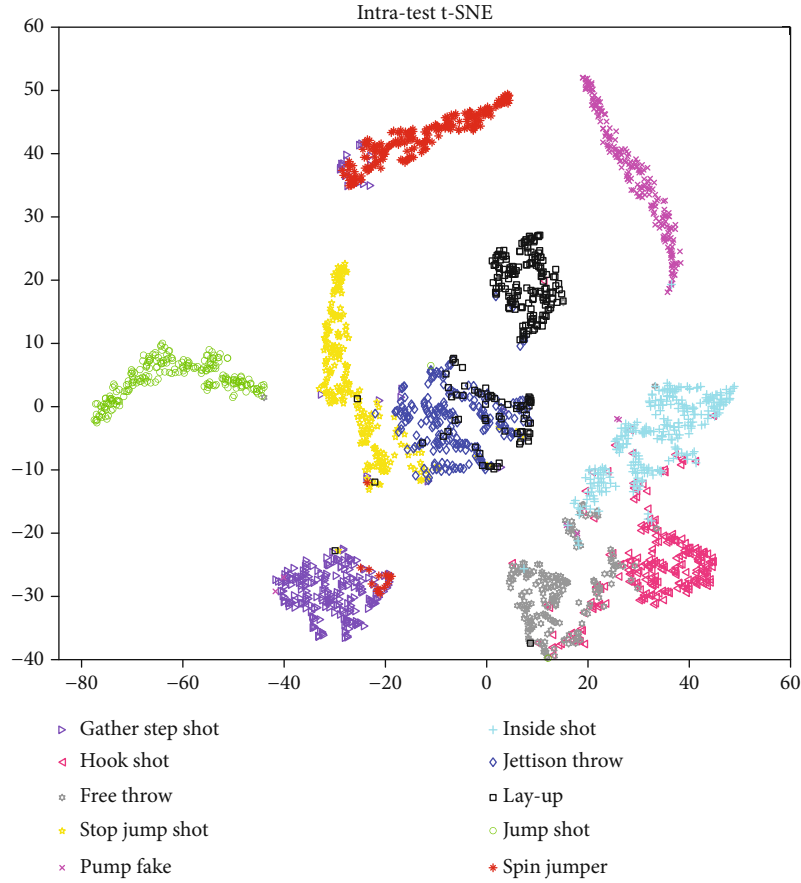


FIGURE 10: t-SNE intratest diagram.

TABLE 4: Classification results of intratest.

Number	Shooting postures	Recall	Precision
1	Gather step shot	96%	99.2%
2	Hook shot	99.2%	100%
3	Free throw	98.1%	98.1%
4	Stop jump shot	98.4%	97.7%
5	Pump fake	98.2%	99.1%
6	Inside shot	99.3%	99.3%
7	Jettison throw	98.8%	98.4%
8	Lay-up	98%	99.2%
9	Jump shot	100%	97.2%
10	Spin jumper	100%	97.9%
	Average	98.6%	98.6%

system still performed well in the intratest, as a result of the selection method of the intratraining and intratest datasets.

Figure 11 presents the t-SNE diagram of the intertest dataset. The t-SNE diagram of intertest dataset demonstrated that inside shot, free throw, hook shot, and jump shot are easily confused, as well as stop jump shot and jettison throw.

Table 5 reports the recall rate, precision rate, average recall rate, and average precision rate of the intertest classification results. The classification results of the intertest show

that the average recall rate was 89.8%, and the maximum recall rate was 100% for the gather step shot, pump fake, inside shot, and spin jumper. The minimum recall rate was 71% for the jettison throw. The average precision rate was 91.1%; the maximum precision rate was 100% for the hook shot and spin jumper, and the minimum precision rate was 65.9% for the inside shot. As per Figure 10, 20 free throws and 74 hook shots were identified as inside shots. This is because free throws and inside shots are similar in action, differing only slightly in the release angle and speed. In addition, as we did not have wireless sensor modules on both left and right wrists, the ability to discriminate between single-handed and double-handed shooting postures is slight. Thus, some of the one-hand shots, such as hook shots, were identified as double-hand shots, such as inside shots. The 35 jettison throws were identified as stop jump shots for the same reason. There are 36 jump shots identified as free throws owing to the similar shooting distance and shooting angle between free throws and jump shots. Furthermore, there are no barometer data collected in this study; thus, there is no clear distinction between jump shooting posture and non-jump shooting posture. In addition, as the subjects in this experiment included five different play positions (i.e., C, PF, SF, SG, and PG), the heights and weights of the subjects were different, the subjects had very different actions in the same shooting posture, and considering that the subjects were

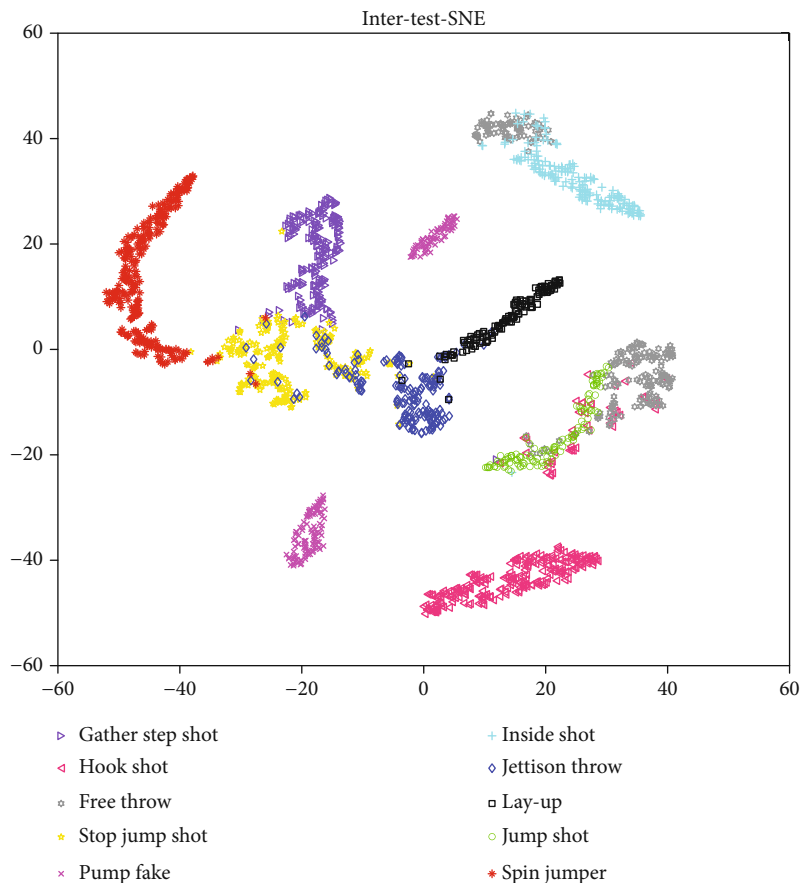


FIGURE 11: t-SNE intertest diagram.

TABLE 5: Classification results of interest.

Number	Shooting postures	Recall	Precision
1	Gather step shot	100%	97.1%
2	Hook shot	72.8%	100%
3	Free throw	89.4%	85%
4	Stop jump shot	94%	86.2%
5	Pump fake	100%	95.8%
6	Inside shot	100%	65.9%
7	Jettison throw	71%	94.2%
8	Lay-up	99.2%	87.4%
9	Jump shot	71.8%	98.9%
10	Spin jumper	100%	100%
	Average	89.8%	91.1%

not professional players, the shooting postures varied considerably. In addition, stability is poor when physical strength is insufficient [34]. These two points also explain the aforementioned low recognition rate of the shooting postures. Finally, the small number of subjects contributes to the low recognition rate.

From the classification results, the VGG16 deep learning model achieved good classification in 10 types of sensor fusion basketball shooting posture recognition experiments.

TABLE 6: Accuracy comparison of classification models.

Literature	[35]	[36]	[37]	This paper
Model	1D-CNN	CNN	FMS-Net	VGG16
Accuracy	90.8%	91.92%	96.7%	98.6%

In contrast, [35] developed a deep learning model around a one-dimensional convolutional network (1D-CNN) architecture and verified it on the public dataset UTD-MHAD, which contained 27 types of activities. In [36], the CNN model was used to identify six types of pedestrian mode. In [37], a hybrid deep learning model based on the fusion of multiple spatiotemporal networks (FMS-Net) was proposed, which was used to detect the four phases of walking. As all the above research results were achieved only for the intratest, the comparison of the above research results with the intratest results of the VGG16 model used in this study found that the classification results of VGG16 were better than those of the other three classification models. The comparison results are shown in Table 6.

To verify the accuracy and effectiveness of the proposed system, it was compared with references [18, 19, 38]. Reference [18] established a real-time wearable assist system for upper extremity throwing action based on accelerometers, which used the longest common subsequence (LCS) algorithm to recognize the six phases of baseball throwing

TABLE 7: Systems comparison.

Literature	Sensor number	Sport	Posture number	Algorithm model	Average accuracy
[18]	2	Baseball	6	LCS	93.9%
[19]	3	Indoor rowing	6	ML	92.4%
[38]	2	Basketball	3	SVM	99.5%
This paper	2	Basketball	10	VGG16	94.3%

posture. In [19], an activity assessment chain for evaluating human activity was established using machine learning (ML) to classify six types of indoor rowing stroke postures (one correct and five incorrect). Reference [38] used a wearable and wireless system based on SVM to recognize overhead passes, chest passes, and shooting in basketball. As shown in Table 7, similar to the three systems above, the system proposed in this paper uses a small number of sensors to recognize a number of postures. This system has certain advantages in terms of average accuracy compared with the systems proposed in references [18, 19]. Although the average accuracy is slightly lower than the system proposed in reference [38], the system proposed in this paper recognizes more postures and achieves good recognition, even for easily confused postures. In addition, compared with the ML and SVM models used in references [19, 38], the deep learning model used in this study has greater development potential. Based on the above analysis, the system proposed in this study exhibits certain advantages compared with the other three systems.

4. Discussion

Shooting is an important aspect in basketball matches and training. Correctly distinguishing the shooting posture used by basketball players in a match and during training can help in making a correct evaluation of the technical characteristics of the players, which in turn could prove helpful in carrying out targeted guidance and practice sessions for players. This study proposes a sensor fusion framework for basketball shooting posture. It fuses the sensor data of the main force hand and main force foot to identify and classify the basic shooting postures and composite shooting postures in basketball. The framework proposed here shows a novel development direction for wearable devices in basketball, which is beyond the conventional framework of IMUs placed only on the arms. Although this sensor fusion framework can recognize more composite shooting postures without integrating more sensors, it still has limitations, which are as follows:

- (1) The amount of limb data is still limited, which can pose challenges for accurately reflecting the subjects' posture information
- (2) While using the proposed method, it is necessary to consider the problem of sensor synchronization and how to align the data of the two sensors if one sensor loses data

- (3) The use of two sensors makes our proposed method more costly compared with the method that uses a single sensor

Many shooting postures in basketball have certain similarities, and nonprofessional basketball fans' shooting postures are generally not standard and less stable; hence, their shooting postures are generally confusing. The basketball shooting posture recognition system proposed in this paper selects nonprofessional basketball fans as subjects, uses a deep learning model based on CNN, classifies 10 types of easily confused shooting postures, and obtains a good classification effect, proving the feasibility of the deep learning model for basketball shooting posture recognition and demonstrating the robustness of the proposed system. Moreover, compared with ML models such as SVM, the deep learning model used in this paper has strong development potential in the future, and it is possible to integrate it into a low-cost integrated circuit in the future to reduce the cost of corresponding smart devices. Therefore, the results of this study can be used in the future for the development of low-cost wearable intelligent basketball motion recognition devices for nonprofessional basketball players.

Basketball shooting, especially composite shooting postures, can be divided into a series of decomposition actions. The time-series and attention of each decomposition action are different [10]. Although the VGG16 deep learning model used in this study has achieved good results in classifying 10 types of sensor fusion basketball shooting postures, there remain shooting postures with lower classification accuracy, such as stop jump shots and inside shots. If time-series and attention judgments are added to the deep learning model, the recognition effect could be further improved. Because researchers have used deep learning models that combine time-series and attention judgments for classification [39], we can also add time-series and attention judgments to deep learning models to improve classification accuracy in our future work. Furthermore, lightweight deep learning models such as MobileNet [40] and SqueezeNet [41] will be adopted to ensure the corresponding time and space efficiency after adding time-series and attention judgments.

Reference [42] studied walking and trotting in equestrian sports by calibrating the sensor data accuracy of four coordinate systems. Similarly, the accuracy of sensor data is also important for recognizing basketball shooting postures. Compared with the reference [42], this study still needs to be strengthened in sensor data accuracy calibration. In general, the sensor will suffer from sensor drift after a period of time, which affects the accuracy of the collected data. In addition, sensor misplacement, as the sensor is not firmly fixed on

the limb, can also lead to other accidents. In [43], the authors proposed a method combining zero velocity update (ZUPT) to reduce the sensor drift error. A rotation matrix method was also proposed in [44] that obtained good performance in dealing with sensor misplacement. In future works, to improve the data precision of sensor fusion for long-term data collection, we will attempt to fix sensor misplacement and sensor drift through software calibration, as was performed in references [42–44]. Furthermore, we will enhance the binding of wireless sensor modules and add a software filter to decrease the effect of sensor misplacement and sensor drift and improve the accuracy of the sensor fusion data.

5. Conclusion

In this study, a sensor fusion basketball shooting posture recognition system based on a CNN was designed. The system used a sensor fusion framework to collect the shooting posture data of the players' main force hand and main force foot and performed sensor data fusion. Subsequently, a CNN-based deep learning model was used for classification. A total of 12,177 sensor fusion basketball shooting posture data entries of the right hand and left foot were collected using this system for 13 Chinese adult male subjects aged 18–40 years with at least 2 years of basketball experience but without any professional basketball training. The shooting posture data entries were trained and tested using the classic VGG16 deep learning model based on CNN through intra- and intertraining/testing methods, achieving satisfactory classification results. These classification results are substantially better than those of similar systems, demonstrating the effectiveness and future development potential of the system.

Data Availability

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

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (Nos. 41971340 and 41271410) and the National Key Research and Development Program of China (Grant no. 2019YFC1510203).

References

- [1] M. Sun, *Advanced Course of Modern Basketball*, Peoples Sports Publishing House, Peking, China, 2018.
- [2] J. Wang, "Ball games," in *Basketball*, pp. 88–120, Higher Education Press, Peking, China, 2009.
- [3] H. Yang, "Essence, characteristics and laws of basketball," *Journal of Chengdu Sport University*, vol. 4, pp. 60–62, 2001.
- [4] J. C. Maglott, J. Xu, and P. B. Shull, "Differences in arm motion timing characteristics for basketball free throw and jump shooting via a body-worn sensorized sleeve," in *Proceedings of the 2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 31–34, Eindhoven, Netherlands, 2017.
- [5] A. Taniguchi, K. Watanabe, and Y. Kurihara, "Measurement and analyze of jump shoot motion in basketball using a 3-D acceleration and gyroscopic sensor," in *Proceedings of the SICE Annual Conference (SICE)*, pp. 361–365, Akita, Japan, 2012.
- [6] M. C. S. Gutiérrez and P. M. V. Castellanos, "Design and validation of a system for improving the effectiveness of basketball players: a biomechanical analysis of the free throw," in *Proceedings of the 2018 IX International Seminar of Biomedical Engineering (SIB)*, pp. 1–8, Bogota, 2018.
- [7] L. Bai, C. Efstratiou, and C. S. Ang, "weSport: utilising wrist-band sensing to detect player activities in basketball games," in *Proceedings of the 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pp. 1–6, Sydney, Australia, 2016.
- [8] Y. Acikmese, B. C. Ustundag, and E. Golubovic, "Towards an artificial training expert system for basketball," in *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, pp. 1300–1304, Bursa, Turkey, 2017.
- [9] L. Zhao and W. Chen, "Detection and recognition of human body posture in motion based on sensor technology," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 15, no. 5, pp. 766–770, 2020.
- [10] D. Hopla, *Basketball Shooting*, The People's Posts and Telecommunications Press, Peking, China, 2020, pp. 17–134.
- [11] S. Shi, Q. F. Zhou, M. Peng, and X. Cheng, "Utilize smart insole to recognize basketball motions," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, pp. 1430–1434, Chengdu, China, 2018.
- [12] M. Peng, Z. Zhang, and Q. Zhou, "Basketball footwork recognition using smart insoles integrated with multiple sensors," in *2020 IEEE/CIC International Conference on Communications in China (ICCC)*, pp. 1202–1207, Xiamen, China, 2020.
- [13] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pp. 253–256, Paris, France, 2010.
- [14] Y. Kim, "Convolutional neural networks for sentence classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1746–1751, Doha, Qatar, 2014.
- [15] J. Gu, Z. Wang, J. Kuen et al., "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018.
- [16] J. Lee, H. Joo, J. Lee, and Y. Chee, "Automatic classification of squat posture using inertial sensors: deep learning approach," *Sensors*, vol. 20, no. 2, p. 361, 2020.
- [17] T. Kautz, B. H. Groh, J. Hannink, U. Jensen, H. Strubberg, and B. M. Eskofier, "Activity recognition in beach volleyball using a deep convolutional neural network," *Data Mining and Knowledge Discovery*, vol. 31, no. 6, pp. 1678–1705, 2017.
- [18] K. Y. Lian, W. H. Hsu, D. Balram, and C. Y. Lee, "A real-time wearable assist system for upper extremity throwing action based on accelerometers," *Sensors*, vol. 20, no. 5, p. 1344, 2020.
- [19] M. Seiffert, F. Holstein, R. Schlosser, and J. Schiller, "Next generation cooperative wearables: generalized activity assessment computed fully distributed within a wireless body area network," *IEEE Access*, vol. 5, pp. 16793–16807, 2017.

- [20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <https://arxiv.org/abs/1409.1556v5>.
- [21] H. Qassim, A. Verma, and D. Feinzimer, "Compressed residual-VGG16 CNN model for big data places image recognition," in *Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 169–175, Las Vegas, NV, USA, 2018.
- [22] Y. Osako, H. Yamane, S.-Y. Lin, P.-A. Chen, and R. Tao, "Cultivar discrimination of litchi fruit images using deep learning," *Scientia Horticulturae*, vol. 269, article 109360, 2020.
- [23] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5325–5334, Boston, Massachusetts, 2015.
- [24] M. B. Priatama, L. Novamizanti, S. Aulia, and E. B. Candrasari, "Hand gesture recognition using discrete wavelet transform and convolutional neural network," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 3, pp. 996–1004, 2020.
- [25] P. Kim, *Design Example Based on MATLAB*, Beihang University Press, Peking, China, 2018.
- [26] H. Lee and M. Whang, "Heart rate estimated from body movements at six degrees of freedom by convolutional neural networks," *Sensors*, vol. 18, no. 5, p. 1392, 2018.
- [27] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification," in *Proceedings of the 2015 IEEE International Conference on Computer Vision*, pp. 1026–1034, Santiago, Chile, 2015.
- [28] L. Z. Zhu, D. M. Chen, J. S. Guo, and H. X. Zhang, "Research on human action recognition based on synergistic LSTM neural network," *Computer Technology Development*, vol. 28, no. 12, pp. 79–82, 2018.
- [29] X. Jiang, B. Hu, S. Chandra Satapathy, S. H. Wang, and Y. D. Zhang, "Fingerspelling identification for Chinese sign language via AlexNet-based transfer learning and Adam optimizer," *Scientific Programming*, vol. 2020, 13 pages, 2020.
- [30] F. Erculj and E. Strumbelj, "Basketball shot types and shot success in different levels of competitive basketball," *Plos One*, vol. 10, no. 6, article e0128885, 2015.
- [31] H. Okubo and M. Hubbard, "Dynamics of the basketball shot with application to the free throw," *Journal of Sports Sciences*, vol. 24, no. 12, pp. 1303–1314, 2006.
- [32] L. Ning, M. Xiao-man, and Z. Ya-hui, "Research status and comments on technical characteristics of single-handed shoulder shooting," *Journal of Guangzhou Sport University*, vol. 39, no. 3, pp. 94–100, 2019.
- [33] J. Krause, D. Meyer, and J. Meyer, *Basketball Skills and Drills*, Posts & Telecom Press, Peking, China, 2017.
- [34] G. Marcolin, N. Camazzola, F. A. Panizzolo, D. Grigoletto, and A. Paoli, "Different intensities of basketball drills affect jump shot accuracy of expert and junior players," *PeerJ*, vol. 6, article e4250, 2018.
- [35] N. Lemieux and R. Noumeir, "A hierarchical learning approach for human action recognition," *Sensors*, vol. 20, no. 17, p. 4946, 2020.
- [36] J. Ye, X. Li, X. Zhang, Q. Zhang, and W. Chen, "Deep learning-based human activity real-time recognition for pedestrian navigation," *Sensors*, vol. 20, no. 9, p. 2574, 2020.
- [37] T. Zhen, L. Yan, and J. L. Kong, "An acceleration based fusion of multiple spatiotemporal networks for gait phase detection," *International Journal of Environmental Research and Public Health*, vol. 17, no. 16, p. 5633, 2020.
- [38] M. Mangiarotti, F. Ferrise, S. Graziosi, F. Tamburrino, and M. Bordegoni, "A wearable device to detect in real-time bimanual gestures of basketball players during training sessions," *Journal of Computing and Information Science in Engineering*, vol. 19, no. 1, 2019.
- [39] X. Li, X. Yi, Z. Liu et al., "Application of novel hybrid deep learning model for cleaner production in a paper industrial wastewater treatment system," *Journal of Cleaner Production*, vol. 294, article 126343, 2021.
- [40] A. Howard, "MobileNets: efficient convolutional neural networks for mobile vision applications," 2017, <https://arxiv.org/abs/1704.04861>.
- [41] F. Iandola, S. Han, M. Moskewicz, K. Ashraf, W. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," 2016, <https://arxiv.org/abs/1602.07360v4>.
- [42] Z. Wang, J. Li, J. Wang et al., "Inertial sensor-based analysis of equestrian sports between beginner and professional riders under different horse gaits," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 11, pp. 2692–2704, 2018.
- [43] S. Qiu, Z. Wang, H. Zhao, and H. Hu, "Using distributed wearable sensors to measure and evaluate human lower limb motions," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 4, pp. 939–950, 2016.
- [44] M. Jiang, H. Shang, Z. Wang, H. Li, and Y. Wang, "A method to deal with installation errors of wearable accelerometers for human activity recognition," *Physiological Measurement*, vol. 32, no. 3, pp. 347–358, 2011.