

## Research Article

# Injury Risk Prediction of Aerobics Athletes Based on Big Data and Computer Vision

Dongdong Zhu,<sup>1</sup> Honglei Zhang,<sup>2</sup> Yulong Sun,<sup>3</sup> and Haijie Qi <sup>4</sup>

<sup>1</sup>*Institute of Physical Education, Dezhou University, Dezhou 253023, Shandong, China*

<sup>2</sup>*Hengshui University, Hengshui 053000, China*

<sup>3</sup>*Hebei North University, Zhangjiakou 075000, Hebei, China*

<sup>4</sup>*Hebei Vocational College of Politics and Law, Shijiazhuang, 050061, China*

Correspondence should be addressed to Haijie Qi; [qihaijie@cumt.edu.cn](mailto:qihaijie@cumt.edu.cn)

Received 21 January 2021; Revised 23 February 2021; Accepted 4 March 2021; Published 2 April 2021

Academic Editor: Shah Nazir

Copyright © 2021 Dongdong Zhu 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, competitive aerobics has been rapidly popularized and developed, and the level of sports skills has also been greatly improved. The performance of some events has gradually approached and reached the advanced level. Therefore, it is vital to invest in the quantitative analysis and cross-disciplinary comprehensive research of aerobics performance and related factors. This paper adopts big data analysis technology and computer vision technology based on convolutional neural network, according to the related theories of sports biomechanics and computer image recognition, to establish a loss risk prediction model for aerobics athletes. The approach firstly has used technology of big data analysis for analyzing the characteristics of competitive aerobics sports data. Secondly, the approach combines the convolutional neural network to visually recognize the aerobics sports images and establish a two-branch prediction model. Finally, the output can be fused to accurately diagnose and evaluate the level of physical fitness development of aerobics athletes, the focus and goal of training content are clarified, and the scientific degree of aerobics training is improved. The study can help injury risk prediction of aerobic athletes based on applications of big data and computer vision.

## 1. Introduction

In recent years, competitive aerobics has developed rapidly in our country, and the corresponding sports injury risks have gradually increased. A number of studies have shown that due to the characteristics of aerobics itself, such as strict time requirements, more difficult movement requirements, fast-paced music accompaniment, and coherent coordinated movements, athletes will suffer sports injuries [1–3] if they are not paying attention. The shoulders, elbows, wrists, waist, thighs, knees, calves, and ankles are the parts that are more prone to injury during aerobics training. Among them, the most prone to injury is the ankle joint [4, 5]. In addition, the types of injury most likely to occur for competitive aerobics athletes is closed injury, most of which are joint strain, sprain and muscle strain, and chronic injuries [6–9], which are the main ones. However, the current scholars'

research on aerobics injuries usually uses questionnaire surveys or expert interviews to determine the injured parts and possible mechanisms of aerobics exercises, and there is a lack of objective empirical research.

In addition, teenagers are in the golden stage of physical development, and various physical qualities will be significantly improved during this period. However, in interviews with young aerobics athletes and coaches, it is found that young athletes have more injuries than adult athletes. This is due to the weakness of the muscles and joints of young athletes, which limits the ability to develop skills and long-term training of irregular technical movements and body postures. In the adolescent period of aerobics athletes, scientific and reasonable training can not only promote the physical development of adolescents but also improve their athletic ability more effectively [10]. Therefore, timely discovery of the causes of injury to young athletes and timely

prevention are critical to improving the skill level of young athletes and extending their sports life.

Due to the characteristics of competitive aerobics, athletes are required to complete a series of high-intensity movements in a short time, which requires a higher level of physical fitness and physical flexibility. Studies have confirmed that long-term high-intensity repetitive exercise training and asymmetric sports skills and postures will increase the risk of athletes' injuries. At the same time, the adolescent stage is a special period of physical development and a period of high incidence of sports injuries [11]. Therefore, this research is based on the development characteristics of adolescents' physical ability and the special characteristics of aerobics. It takes college students in a certain city as the research object, and the tests are conducted in groups of different genders and aerobics sports grades, and uses the receiver operating characteristic (ROC) curve to formulate evaluation standards, analyze the physical weakness of young aerobics athletes, and evaluate the risk of noncontact injury to provide a certain theoretical reference for subsequent aerobics training for young people. This paper presents a method of motion recognition for aerobics athletes based on machine vision, which recognizes joint strain, sprain, and muscle strain caused by their movements, and uses big data to train the vision-based deep learning [12–16] algorithm of this article. Following are the main innovative points of this paper:

- (i) This article proposes a method of motion recognition for aerobics athletes based on machine vision, which recognizes joint strain, sprain, and muscle strain caused by their movements.
- (ii) This paper constructs a dual-branch injury risk prediction model for aerobics athletes. One branch uses big data to analyze the characteristics of aerobics athletes' movement injuries, and the other branch builds a deep convolutional neural network model to identify joint strain, sprain, and muscle strain and perform counting and prediction.
- (iii) We have conducted sufficient comparative experiments and ablation studies to prove the effectiveness of the algorithm based on big data and computer vision proposed in this paper. It can be used to discover the causes of injury to young athletes in time and prevent them in time, which is useful for improving the skills of young athletes, and prolonging sports life.

The organization of the paper is as follows: Section 2 shows the related research to the proposed area of research. Section 3 represents the methodology section of the proposed study with details of the given approach. Section 4 shows the experiments and results of the current study. The paper is concluded in Section 5.

## 2. Related Research

Various approaches and techniques have been devised in the literature for injury risk prediction of aerobics athletes.

Fanian et al. [17] studied 6755 cases of sports injuries and investigated and found that the sport with the highest incidence is football; the common injury sites are knee joints (15%), calves (11.9%), and wrist joints (11%). The most common types of injuries are fractures (12.8%), cartilage injuries (6.87%), and contusions (2.5%); 12% of the patients underwent surgery; the average hospital stay was 3–4 days. Michael [18] surveyed athletes in various sports such as gymnastics, basketball, football, and running. After summarizing and analyzing the types and characteristics of athletes' sports injuries in different sports, they proposed the diagnosis and treatment of various types of sports injuries. Method aims at the occurrence of injury and points out two major steps to reduce the occurrence of sports injuries: one is to wear protective gear in vulnerable parts to reduce the probability of injury and the degree of injury; the other is to improve the athletes as much as possible. Physical fitness, including strength, agility, and flexibility, reduces the chance of athletes' injuries. Wiese-Bjornstal et al. [19] pointed out that sports injury is a relatively unacceptable thing for athletes. After a sports injury occurs, athletes will have complex psychology during training and rehabilitation, which is mainly reflected in cognition, emotion, and behavior. On the one hand, when athletes experience the process of recovering from injury and returning to training, they often have certain cognitive and emotional reactions, which are mainly affected by the individual and the environment. In addition, the time of injury will also affect the psychology of injured athletes to varying degrees. When sports injuries occur shortly before major competitions, the athletes' sense of disappointment and despair will be greatly enhanced [20].

Malliou et al. [21] used questionnaire surveys and conducted statistics and analysis on the injuries of athletes engaged in aerobics training and found that, in the injured population, lower limb injuries accounted for 97.3% and ankle and knee injuries were the most common in aerobics training. At the same time, it is pointed out that training time, years, and training level will have an impact on injury. Bintoudi et al. [22] investigated two aerobics pedal athletes and found that they had knee joint pain and fat pad edema. The article discussed the possible pathogenic factors and mechanisms of aerobics. Kiesel et al. [23] found in a study of professional American rugby players during nonseason that the scores of the FMS test are directly proportional to the athlete's sports injury risk. Through the calculation of the receiver operating curve, a benchmark score of 14 points is delineated for rugby players. Athletes with a season average FMS score of less than 14 points have a much higher risk of injury than those with more than 14 points. In his subsequent similar research, he found that after targeted intervention training for athletes, the number of FMS scores greater than 14 increased significantly, and athletes with movement asymmetry problems could also be improved. Dennis Rex [24] conducted FMS tests on 67 college football players and combined their lower limb explosive power and season injury data. They believed that the FMS test can be used as an assessment tool to predict serious sports injuries and reminded that the FMS score was less than 11 points.

The sports injury risk of athletes is 9 times that of other athletes, and timely corrective training is required. Dorrel et al. [25] followed up on 257 college athletes' injuries after conducting FMS tests and found that those athletes who scored less than 15 on the FMS test were more likely to get injured.

### 3. Methodology

This section mainly introduces the big data platform technology and related theories related to the research of this article. Figure 1 shows the examples of aerobics sports injuries.

The proposed approach firstly introduces the core architecture, working principle, and related ecosystem of the Hadoop platform and then introduces the basic ideas of the data mining theory of big data analysis and the common algorithms of data mining are outlined. Finally, the principle of convolutional neural network [26–31] and the two-branch algorithm proposed in this paper are explained. Figure 2 shows the overall framework of the proposed approach.

**3.1. Data Mining and Analysis.** The following subsections show the data mining and analysis section of the paper.

**3.1.1. Big Data Platform.** Hadoop is currently one of the big data platforms widely used by relevant research institutions in the industry. It is an open-source top-level project maintained by the Apache Foundation. Hadoop inherits the idea of distribution and fully applies it to data storage and processing technology. The Hadoop ecosystem is also composed of many open source software. The open source of the entire ecosystem allows the software in it to be supervised and maintained by everyone, and the stability and security of the system will have a relatively high guarantee. Moreover, the open source of the entire system greatly facilitates the use of users. Big data computing is no longer the unique ability of some professional organizations and has gradually become a field that every developer can set foot in. Figure 3 is its ecosystem architecture.

HDFS is the distributed file system of Hadoop. As the data foundation of the entire ecosystem, HDFS is located at the bottom of the ecosystem. It has very low requirements for hardware resources, can run on computers with low configuration, and can ensure the safety and stability of data in the event of hardware failure through redundant storage. The file system divides each file into multiple parts, each part is called a data block, and each data block is copied into three copies and stored in different locations, which not only provides high throughput of data access but it also has extremely high fault tolerance. Figure 4 shows the HDFS architecture diagram. As the role of the manager in HDFS, NameNode undertakes the core tasks. It is responsible for managing the mapping information of data blocks. As the hot backup storage of the NameNode, the significance of the existence of the secondary and the NameNode is to cope with the single point of failure that may occur in the NameNode. HDFSClient is a client that accesses HDFS, and

all client requests will first interact with the NameNode. Each DataNode maintains interaction with the NameNode and serves as a block where the slave node stores data.

MapReduce is the computing engine of Hadoop. JobTracker is the core of MapReduce, responsible for all task allocations and job scheduling. JobTracker decomposes tasks and distributes them to each TaskTracker node for execution. TaskTracker runs maps and reduce tasks and reports the status of tasks to JobTracker regularly. There are many important components in the Hadoop ecosystem. HBase is a high-performance, scalable column storage database for structured data. Hive originated from Facebook, which is a Hadoop-based data warehouse, mainly used to solve the problem of massive structured data statistics. The biggest feature of Hive is that it can convert SQL into MapReduce jobs so that it can be executed on Hadoop. Zookeeper is used to solve the problem of application coordination in a distributed environment. Sqoop is responsible for data transmission tasks between traditional databases and Hadoop. Pig can convert scripting languages into MapReduce jobs and execute them on Hadoop. Mahout is a library of machine learning algorithms for Hadoop. Flume is an open source log collection system.

**3.1.2. Data Mining.** As a data processing technology, data mining aims to discover the laws and knowledge behind the phenomena hidden in reality through a series of operations and calculations on data. It is also one of the current research hotspots in the scientific field. Data mining technology draws on the advantages of a variety of related technologies, which can extract hidden valuable information from real data and provide references for actual activities. From the methodological point of view, data mining can be divided into two categories: description and prediction. The similarity between the two is that the law is calculated through the existing data. The difference is that the purpose of the description is to provide interpretation support for the law of the data, and the prediction is to provide forecasts for actual activities.

The functions of data mining are divided into the following categories:

(1) Concept or class description:

This kind of data mining mainly describes the class or concept of data through the method of data differentiation and characterization. Data classification classifies the data to be mined by constructing a comparative dataset. Data characterization is by first querying existing related datasets and then summarizing their characteristics.

(2) Predictive modeling:

Data mining prediction methods are divided into two categories, among which classification and regression methods are applied to discrete and continuous variables. Predictive modeling derives a model through training so that the error between the predicted value and the actual value of the specified variable reaches the global minimum. The

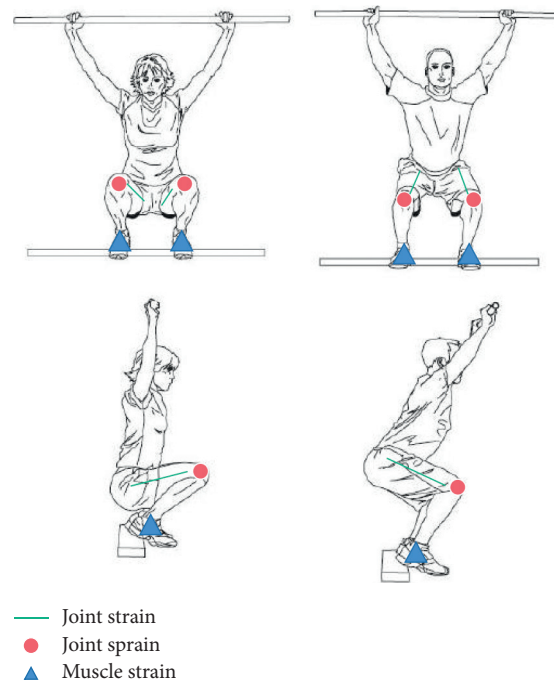


FIGURE 1: Examples of aerobics sports injuries.

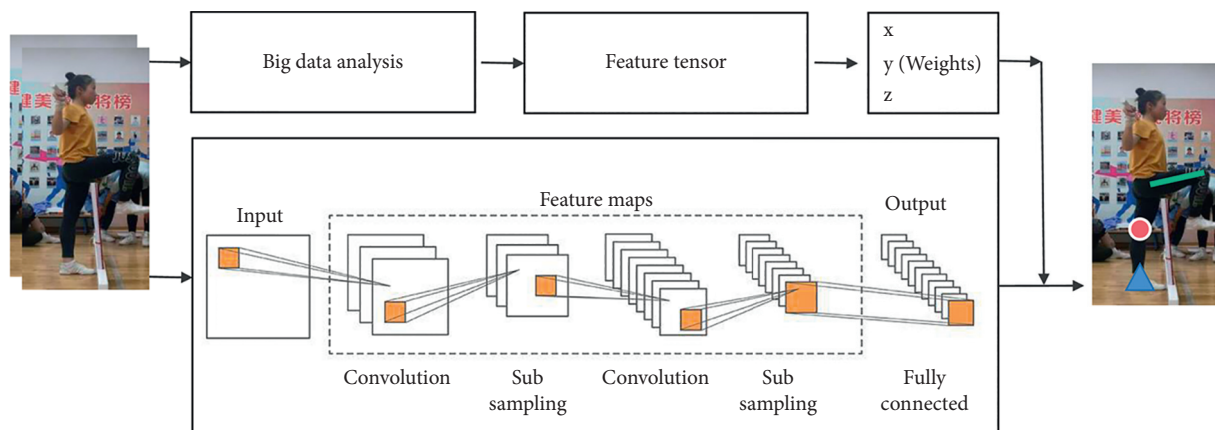


FIGURE 2: Overall framework of our proposed algorithm.

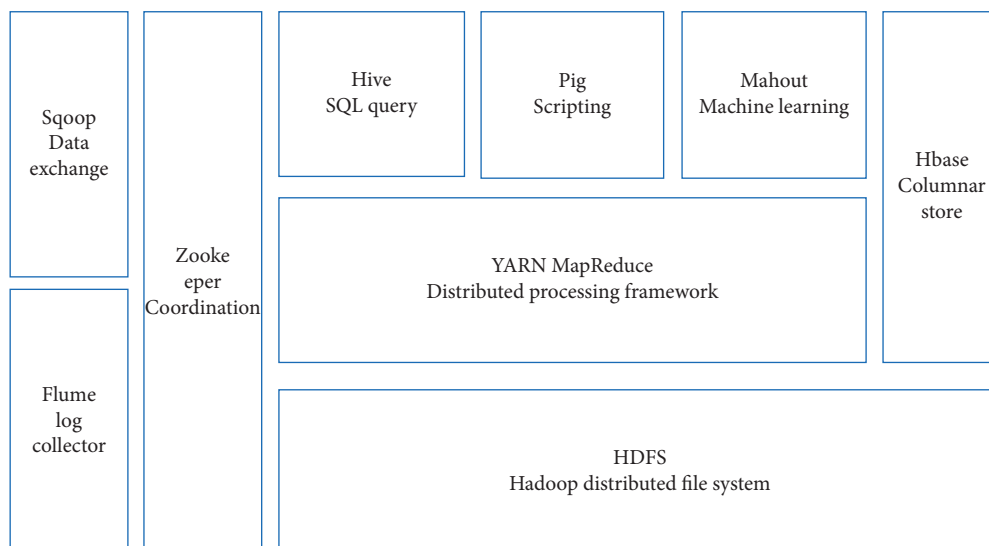


FIGURE 3: Hadoop ecosystem architecture.



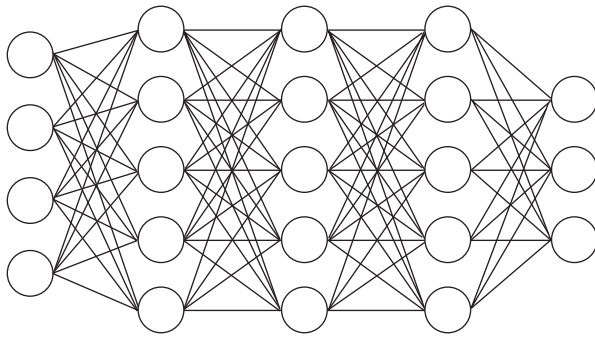


FIGURE 4: Schematic diagram of fully connected neural network.

application of this type of algorithm is commonly used in disease risk prediction.

(3) Association analysis:

Used to describe the associated features in the data, the discovered patterns are usually expressed in the form of implicit rules or feature subsets. The association rules derived from the association analysis can reveal the dependence of each element in the dataset and indicate the conditions under which the attributes appear in the dataset with a specific frequency. For example, the correlation analysis of multiple symptoms of patients can find the correlation rules that occur between the symptoms.

(4) Cluster analysis:

The main function of this type of data mining algorithm is to divide the dataset into valuable or meaningful groups. Cluster analysis and classification analysis are both similar and different. The difference is that cluster analysis belongs to unsupervised classification. The class label needs to be obtained from the data, and the number of classes is not known before clustering.

(5) Abnormal detection:

Based on the analysis of the dataset, abnormal characteristics or abnormal data are obtained. It is often used in the diagnosis of abnormal diseases and the detection of abnormal network traffic.

**3.1.3. Data Mining Algorithm.** As data mining continues to incorporate new domain knowledge, data mining algorithms are constantly evolving and mining methods are becoming more abundant. People can choose specific mining algorithms based on mining needs. The following are several commonly used algorithms for data mining:

(1) Neural network:

Neural network is an intelligent algorithm that simulates human brain nerve transmission. It is generally composed of three parts: input, output, and implicit. It is divided into three models. They are feedforward network model, feedback network model, and self-organization network model. The neural network has strong ability to process

nonlinear data, good fault tolerance performance, and high classification accuracy, but low performance in the interpretation of the results. The most commonly used in the medical field is a multilayer feedforward neural network, the BP neural network. It is worth noting that the neural network used in this article is a convolutional neural network.

(2) Decision tree:

Decision tree is a tree structure model with classification rules, and its logical branch relationship is top-down. The decision tree selects the root node according to the variable attributes according to parameters such as information gain and Gini coefficient and then divides down according to the variable attributes of the root node to form branches; then, each branch node retests the variable attributes and continues to branch down and so on continue until the node's category is homogenized or reaches the set threshold. The algorithm can be converted into classification rules to classify diseases according to their symptoms, thereby predicting diseases. Common algorithms include ID3 algorithm, C4.5 algorithm, and CART algorithm.

(3) Clustering algorithm:

The purpose of clustering algorithm is to obtain several classes through computational analysis, which is an unsupervised learning method. Among them, the data between different classes are usually uncorrelated or there are certain differences, and the data between the same classes have a certain correlation or similarity. The k-means algorithm is the most common clustering algorithm.

(4) Association rules:

Association rule mining is the process of finding strong association rules through the specified minimum support and minimum confidence. Usually, it consists of two parts, one is to find all frequent item sets, and the other is to find the association rules in frequent item sets. Common algorithms include Apriori algorithm and FP-growth algorithm.

(5) Association classification:

Association classification algorithm is one of the important classification methods. The characteristic of this classification method is to first extract the association classification rules and then build a model to predict unknown instances and combine association rule mining and classification. Association classification algorithms usually consist of three parts: rule generation, rule sorting and pruning, and prediction of new instances. Common algorithms include CBA algorithm, CMAR algorithm, and ACSER algorithm.

**3.2. Convolutional Neural Networks.** Figure 4 is a structural diagram of a fully connected neural network, and Figure 5 is a structural diagram of a convolutional neural network.

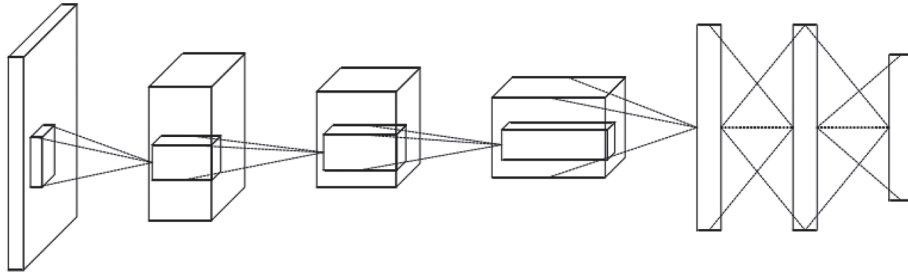


FIGURE 5: Schematic diagram of convolutional neural network.

Although the two are quite different on the surface, they are actually very similar in structure. Convolutional neural networks are also connected through layers of nodes, and each node also represents a neuron. The difference is that a fully connected neural network usually has a connection to every neuron in every two adjacent layers, while only some nodes in the adjacent layers of a convolutional neural network are connected, which can effectively alleviate the problem of excessive parameters of the fully connected neural network. Too many parameters of the neural network will cause the calculation speed to be slow, the calculation time is too long, and at the same time, it is more prone to overfitting. Convolutional neural networks can effectively reduce the number of parameters and speed up model training.

In the first few layers of the convolutional neural network, the data nodes are transformed into a three-dimensional matrix, and only some nodes are connected to adjacent layers. A convolutional neural network usually consists of the following structure.

**3.2.1. Input Layer.** The input layer is the input of the neural network. In the image-oriented convolutional neural network, it represents the pixel matrix converted after a picture is read by a computer. The pixel matrix is a three-dimensional matrix. The length and width represent the size of the image, and the depth represents the number of color channels of the image. When the input picture is a black and white picture, the depth is 1; when the input picture is a color picture, the depth is 3. Starting from the input layer, the three-dimensional matrix is transformed into another three-dimensional matrix through a different network structure until the final fully connected layer.

**3.2.2. Convolutional Layer.** The convolutional layer is the most important part of the convolutional neural network and the key to extracting features.

Figure 6 is a schematic diagram of the convolutional layer transformation. The small matrix of the input layer and the convolution kernel are convolved to obtain the small matrix of the output layer. The convolution kernel is also a three-dimensional matrix. Its length and width are manually set. The convolution kernel is also a three-dimensional matrix, its length and width are manually set, and the size is  $3 \times 3$  or  $5 \times 5$ . Since the depth of the two matrices of the convolution operation must be the same, the depth of the

filter cannot be changed and must be the same as the depth of the input layer matrix. The convolution kernel also needs to manually set the number, and its number determines the depth of the output layer matrix. The process from the input layer to the output layer is called forward propagation. Assuming that  $a_{x,y,z}$  is used to represent the value of a certain point in the input matrix,  $w_{x,y,z}$  represents the value of the position of the convolution kernel, and  $b$  represents the node's corresponding bias term parameter, then

$$g = f \left( \sum_{x=1}^m \sum_{y=1}^n \sum_{z=1}^q a_{x,y,z} \times w_{x,y,z} + b \right), \quad (1)$$

where  $g$  is the output of the corresponding point,  $(m, n, q)$  are the length, width, and number of the convolution kernel, respectively, and  $f$  is the activation function. The activation function acts as a nonlinear transformation of the output. The commonly used activation function consists of a linear rectification unit (ReLU) and a hyperbolic function (sigmoid).

**3.2.3. Pooling Layer.** It can be seen from Figure 7 that a pooling layer is usually added after the convolutional layer. The pooling layer can reduce the size of the input data, reduce the number of parameters, and speed up the network calculation at the same time. Preventing the occurrence of overfitting problems is similar to the forward propagation process of the convolutional layer. The pooling layer is also completed by sliding a filter structure on the matrix. The difference is that the pooling layer uses a simple maximum value or the average method.

Figure 7 shows the calculation process of the pooling layer using maximum pooling. Here, the size of the filter is  $2 \times 2$ , that is, the target area of a pooling operation is  $2 \times 2$  nodes, and the step size is 2, that is, the calculation is done in every 2 steps.

**3.2.4. Fully Connected Layer.** The fully connected layer of CNN, like the fully connected neural network, maps the features learned by the previous convolutional layer and the pooling layer to the sample label space, which acts like a classifier. The fully connected layer is not necessary; it can be replaced by a convolutional layer using a convolution kernel of size  $1 \times 1$ .

**3.2.5. Softmax Layer.** In the use of multiclassification, the softmax layer is often used as the last layer of the deep neural network, which makes the output of the network straightforward to the probabilities of various classifications. The three aerobics sports injury categories in this article are multiclassification tasks. where  $n$  is the number of training samples,  $y$  is the label value of the training data,  $L$  is the number of network layers, and  $a^L(x)$  is the output of the network.

Introduce the mean square error function as the cost function in back propagation,

$$C = \frac{1}{2n} \sum_x \|y - a^L(x)\|^2, \quad (2)$$

Calculate the error of each unit of the output layer, the definition of the error of the  $j$ th unit of the  $L$ -th layer:

$$\delta_j^L = \frac{\partial C}{\partial z_j^L}. \quad (3)$$

According to the chain rule, we have

$$\delta_j^L = \frac{\partial C}{\partial z_j^L} = \frac{\partial C}{\partial a_j^L} \frac{\partial a_j^L}{\partial z_j^L} = \frac{\partial C}{\partial a_j^L} f'(z_j^L). \quad (4)$$

Its matrix form is

$$\delta^L = \left( (w^{L+1})^T \delta^{L+1} \right) \otimes f'(z^L). \quad (5)$$

After calculating the error of the output layer, the error of the previous layer should be calculated. According to the chain rule, we can get

$$\begin{aligned} \delta_j^l &= \frac{\partial C}{\partial z_j^l} \\ &= \sum_k \frac{\partial C}{\partial z_k^{l+1}} \frac{\partial z_k^{l+1}}{\partial z_j^l} \\ &= \sum_k \frac{\partial z_k^{l+1}}{\partial z_j^l} \delta_k^{l+1}. \end{aligned} \quad (6)$$

Therefore,

$$\delta_j^l = \sum_k w_{kj}^{l+1} \delta_k^{l+1} f'(z_j^l). \quad (7)$$

In this way, errors can run through the entire network by back propagation. Let joint strain, sprain, and muscle strain be  $x$ ,  $y$ , and  $z$ . The output equation through the softmax function is

$$\text{Soft max}(\text{input\_image}) = \max[x, y, z]. \quad (8)$$

## 4. Experiments and Results

The following subsections show the experiments and results of the paper.

### 4.1. Experimental Environment and Hyperparameter Settings.

This experiment uses the PyCharm compiler and the TensorFlow deep learning framework in the Windows environment and uses the small batch learning method, the Adam optimizer, the learning rate is 0.0001, 100 images are selected for training at a time, and a total of 1000 iterations are performed. The experimental model parameters are shown in Table 1.

**4.2. Subjects of the Experiment.** Taking the freshman and sophomore aerobics students in a certain city as the research object, there are a total of 60 girls. At the beginning of the experiment, there were 2 people with acute sports injuries, and they were finally confirmed as 58 people. There was no significant difference in the height, weight, age, and training years of the 58 aerobics professional girls.

**4.3. Experimental Methods.** The visual recognition experiment based on convolutional neural network in this paper is divided into training phase and testing phase. The training phase has the following contents.

**4.3.1. Aerobics Squat (Deep Squat).** Stand upright with your feet shoulder-width apart or slightly wider than shoulder-width apart, with your toes pointing straight ahead. Place the test rod on the top of the head and adjust your hands to make the elbow joint 90°; extend both arms at the same time to lift the test rod to the head. Then, slowly squat until your thighs are parallel to the ground, keep your heels from leaving the ground, raise your head and chest, keep your back straight, do not arch your waist, and hold the test rod as far above your head as possible. During the squat process, the knee joints on both sides are in the same plane as the feet. Do not buckle your knees inward, and always keep your toes pointing forward. This squat test was repeated three times. This experiment takes photos of the entire training process for the neural network to learn the characteristics of joint strain, sprain, and muscle strain.

**4.3.2. Active Straight Leg Raise.** The subject lies on his back, with his arms on his side, palms facing up, and lying flat; the test board is placed under the knees with the toes pointing directly up.

It is placed in the middle of the line between the anterior superior iliac spine and the midpoint of the patella, perpendicular to the ground. The subject lifted one leg to the maximum extent and kept the knee joint straight during the lifting process. The knee joint on the other side should be kept in contact with the test board as much as possible, with the toes pointing upward, and the opposite side was also the same. This experiment takes photos of the entire training process for the neural network to learn the characteristics of joint strain, sprain, and muscle strain.

The training process is 1 week, and the test actions are 50 sets. The experimental results are shown in Table 2.

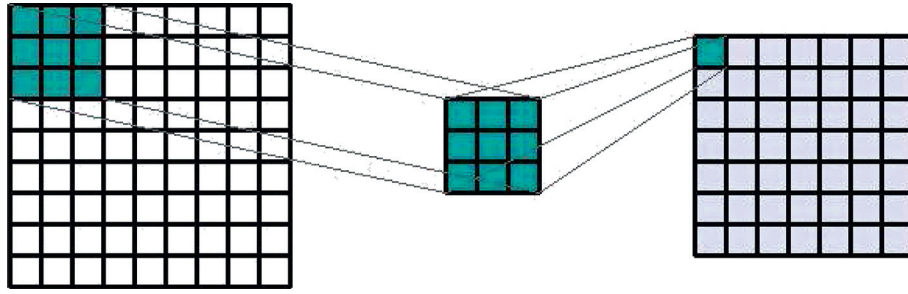


FIGURE 6: Convolutional layer.

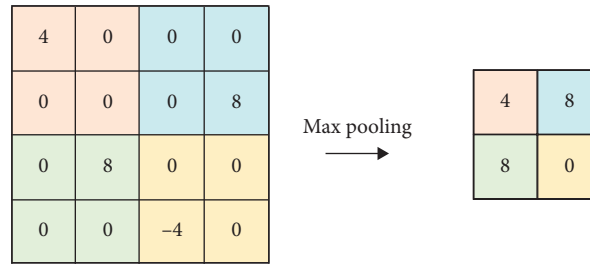


FIGURE 7: Max pooling layer.

TABLE 1: Experimental environment configuration.

	Filter size	Number of filters	Sliding step	Zero padding
Conv layer	5 * 5	32	2/1	2
Pooling layer	2 * 2		2	0
Conv layer	3 * 3	64	1	1
Pooling layer	2 * 2		2	0
Conv layer	3 * 3	128	1	1
Pooling layer	2 * 2		2	1
FC layer		512		
FC layer		256		
FC layer		128		
FC layer		100		
Softmax layer		3		

TABLE 2: Aerobics sports injury prediction results.

	Joint strain (%)	Joint sprain (%)	Muscle strain (%)
Group 1	74	68	71
Group 2	69	64	77
Group 3	78	64	69
Group 4	79	79	71
Group 5	80	61	69
Group 6	74	71	79

The CNN training results are shown in Figure 8. When the number of training steps is greater than 900, the loss function value drops below 0.1.

**4.4. Visualization of Results.** It can be seen from Figure 9 that the acute sports injury rate of aerobics in the low-risk group

is lower than that of the high-risk group. Therefore, regardless of whether there is a corrective training intervention, the injury rate of the two low-risk groups will not be very high. The members of the high-risk group have a higher risk of injury during the training process. Therefore, the injuries of the two high-risk groups can better verify the effectiveness of the corrective training. This experimental



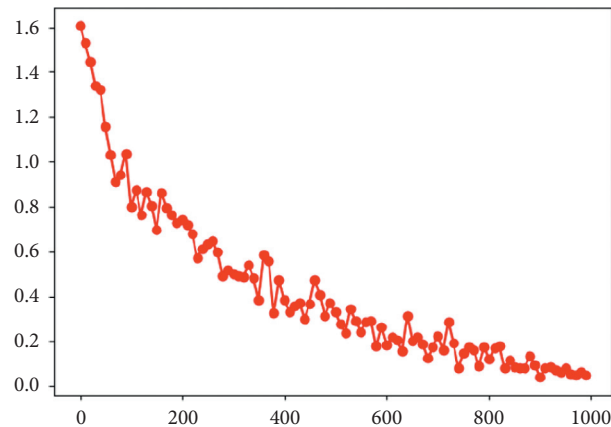


FIGURE 8: Declining curve of loss function.

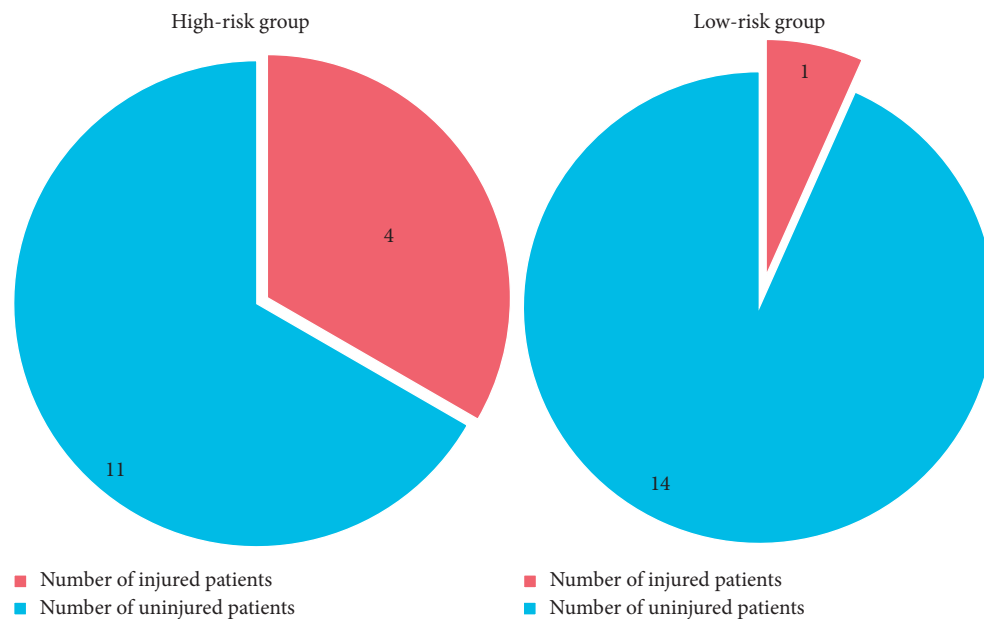


FIGURE 9: Comparison results of injury rate prediction between high-risk group and low-risk group.

test proves the effectiveness of the algorithm in this paper and can improve the scientific degree of aerobics training.

## 5. Conclusion

Competitive aerobics has been rapidly promoted and established, and the level of sports expertise has also been significantly enhanced. The performance of some events has progressively approached and reached the advanced level. Therefore, it is vital to invest in the quantitative analysis and cross-disciplinary wide-ranging research of aerobics performance and associated factors. This paper constructs a novel dual-branch aerobics athlete injury risk prediction algorithm based on big data and computer vision technology, and through experimental research, it has been proved that big data analysis can extract effective features from competitive aerobics data. Secondly, combined with a convolutional neural network to visually recognize aerobics images, it can accurately diagnose

and evaluate the physical fitness development level of aerobics athletes, clarify the focus and objectives of the training content, and improve the scientific degree of aerobics training.

## Data Availability

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

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

- [1] C. Bolling, W. Van Mechelen, H. R. Pasman, and E. Verhagen, "Context matters: revisiting the first step of the 'sequence of

- prevention' of sports injuries," *Sports Medicine*, vol. 48, no. 10, pp. 2227–2234, 2018.
- [2] A. Carbone and S. Rodeo, "Review of current understanding of post-traumatic osteoarthritis resulting from sports injuries," *Journal of Orthopaedic Research*, vol. 35, no. 3, pp. 397–405, 2017.
  - [3] A. Olmedilla-Zafra, V. J. Rubio, E. Ortega, and A. García-Mas, "Effectiveness of a stress management pilot program aimed at reducing the incidence of sports injuries in young football (soccer) players," *Physical Therapy in Sport*, vol. 24, pp. 53–59, 2017.
  - [4] X. Xiao, W. Xiao, X. Li, B. Wan, and G. Shan, "The influence of landing mat composition on ankle injury risk during a gymnastic landing: a biomechanical quantification," *Acta of Bioengineering and Biomechanics*, vol. 19, no. 1, 2017.
  - [5] J. R. Miller, K. W. Dunn, L. J. Ciliberti Jr., S. W. Eldridge, and L. D. Reed, "Diagnostic value of early magnetic resonance imaging after acute lateral ankle injury," *The Journal of Foot and Ankle Surgery*, vol. 56, no. 6, pp. 1143–1146, 2017.
  - [6] S. Malone, M. Roe, D. A. Doran, T. J. Gabbett, and K. Collins, "High chronic training loads and exposure to bouts of maximal velocity running reduce injury risk in elite Gaelic football," *Journal of Science and Medicine in Sport*, vol. 20, no. 3, pp. 250–254, 2017.
  - [7] X. Deng, X. Zhang, W. Li et al., "Chronic liver injury induces conversion of biliary epithelial cells into hepatocytes," *Cell Stem Cell*, vol. 23, no. 1, pp. 114–122, 2018.
  - [8] Y. A. Tuakli-Wosornu, E. Mashkovskiy, T. Ottesen, M. Gentry, D. Jensen, and N. Webborn, "Acute and chronic musculoskeletal injury in para sport," *Physical Medicine and Rehabilitation Clinics of North America*, vol. 29, no. 2, pp. 205–243, 2018.
  - [9] L. Bowen, A. S. Gross, M. Gimpel, and F.-X. Li, "Accumulated workloads and the acute:chronic workload ratio relate to injury risk in elite youth football players," *British Journal of Sports Medicine*, vol. 51, no. 5, pp. 452–459, 2017.
  - [10] S. J. Biddle, C. J. Wang, N. L. Chatzisarantis, and C. M. Spray, "Motivation for physical activity in young people: entity and incremental beliefs about athletic ability," *Journal of Sports Sciences*, vol. 21, no. 12, pp. 973–989, 2003.
  - [11] W. Van Mechelen, H. Hlobil, and H. C. G. Kemper, "Incidence, severity, aetiology and prevention of sports injuries," *Sports Medicine*, vol. 14, no. 2, pp. 82–99, 1992.
  - [12] X. Ning, K. Gong, W. Li, L. Zhang, X. Bai, and S. Tian, "Feature refinement and filter network for person Re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, In press.
  - [13] W. Cai and Z. Wei, "PiiGAN: generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.
  - [14] X. Ning, P. Duan, W. Li, and S. Zhang, "Real-time 3D face alignment using an encoder-decoder network with an efficient deconvolution layer," *IEEE Signal Processing Letters*, vol. 27, pp. 1944–1948, 2020.
  - [15] W. Cai and Z. Wei, "Remote sensing image classification based on a cross-attention mechanism and graph convolution," *IEEE Geoscience and Remote Sensing Letters*, 2020.
  - [16] X. Ning, K. Gong, W. Li, and L. Zhang, "JWSAA: joint weak saliency and attention aware for person re-identification," *Neurocomputing*, 2020, in Press.
  - [17] H. Fanian, M. Khalilian, A. Zamani et al., *XXIX FIMS World Congresses of Sports Medicine*, pp. 14–16, Beiting, Sanya, China, 2006.
  - [18] L. Michael, Tuggy & Cora Collette Brenner. *Taylor Musculoskeletal Problems and Injuries a Handbook*, pp. 205–231, Springer, New York, NY, USA, 2006.
  - [19] D. M. Wiese-Bjomstal, A. M. Smith, S. M. Shaffer et al., "An intergraded model of response to sport injury: psychological and societal dynamics," *Journal of Applied Sport Psychology*, vol. 10, no. 10, pp. 46–69, 1993.
  - [20] T. Bianco, S. Malo, and T. Orlick, "Sport injury and illness: elite skiers describe their experiences," *Research Quarterly for Exercise and Sport*, vol. 70, no. 2, pp. 157–169, 1999.
  - [21] P. Malliou, A. Gioftsidou, G. Pafis, A. Beneka, and G. Godolias, "Proprioceptive training (balance exercises) reduces lower extremity injuries in young soccer players," *Journal of Back and Musculoskeletal Rehabilitation*, vol. 17, no. 3-4, pp. 101–104, 2004.
  - [22] A. Bintoudi, M. Goumenakis, and A. Karantanis, "Suprapatellar fat pad inflammation in step aerobics athletes: MR imaging evaluation of two cases," *Open Medicine*, vol. 7, no. 6, pp. 813–816, 2012.
  - [23] K. Kiesel, P. J. Plisky, and M. L. Voight, "Can serious injury in professional football be predicted by a preseason functional movement screen," *North American Journal of Sports Physical Therapy*, vol. 2, no. 3, pp. 147–158, 2007.
  - [24] B. Dennis Rex, *Strength Flexibility Functional Movement and Injury in Collegiate Men Football Players*, Georgia College and State University, Milledgeville, Georgia, 2010.
  - [25] B. Dorrel, T. Long, S. Shaffer, and G. D. Myer, "The functional movement screen as a predictor of injury in national collegiate athletic association division II athletes," *Journal of Athletic Training*, vol. 53, no. 1, pp. 29–34, 2018.
  - [26] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, pp. 1–22, 2021, In press.
  - [27] X. Ning, W. Li, B. Tang, and H. He, "BULDP: biomimetic uncorrelated locality discriminant projection for feature extraction in face recognition," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2575–2586, 2018.
  - [28] Z. Wang, C. Zou, and W. Cai, "Small sample classification of hyperspectral remote sensing images based on sequential joint deeping learning model," *IEEE Access*, vol. 8, pp. 71353–71363, 2020.
  - [29] W. Cai and Z. Wei: (2019). Diversity-Generated Image Inpainting with Style Extraction. <https://arxiv.org/pdf/1912.01834.pdf>.
  - [30] Z. L. Yang, S. Y. Zhang, Y. T. Hu, Z. W. Hu, and Y. F. Huang, "VAE-Stega: linguistic steganography based on variational auto-encoder," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 880–895, 2020.
  - [31] X. Ning, X. Wang, S Xu et al., *A Review of Research on Co-Training. Concurrency and Computation: Practice and Experience*, Wiley, Hoboken, NJ, USA, 2021, In press.