

## Research Article

# Sports Injury Rehabilitation Intervention Algorithm Based on Visual Analysis Technology

Xiao Chen<sup>1</sup> and Guoliang Yuan <sup>2</sup>

<sup>1</sup>Physical Education Department, Zhongnan University of Economics and Law, Wuhan 430073, Hubei, China

<sup>2</sup>College of Physical Education, Hengshui University, Hengshui 053000, Hebei, China

Correspondence should be addressed to Guoliang Yuan; [yanguoliang053000@163.com](mailto:yanguoliang053000@163.com)

Received 28 March 2021; Revised 9 April 2021; Accepted 19 April 2021; Published 24 May 2021

Academic Editor: Fazlullah Khan

Copyright © 2021 Xiao Chen and Guoliang Yuan. 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.

Sports injuries of high-level athletes restrict the improvement of sports performance. Under this premise, an efficient and accurate sports injury assessment method is needed to detect potential sports injuries and conduct injury prevention training. Therefore, this paper proposes a novel sports injury prediction algorithm based on visual analysis technology. The proposed algorithm first takes the time-frequency of sensed data as the convolutional neural network (CNN) input. The one-dimensional time series collected by the sensor is converted into two-dimensional images using the Gram angle domain algorithm. The one-dimensional sensed data provides a new perspective and provides a basis for better use of convolutional neural networks and computer vision technology. Second, combining the residual network's structure and advantages and hole convolution and multihole convolution kernel residual module is proposed. It improves the model's ability to extract features at different scales while effectively controlling the parameter scale. Based on these modules, a single-sensor-based athlete action recognition algorithm is proposed. Several comparative experiments have been conducted on a public data set containing only acceleration sensors to verify the proposed algorithm's effectiveness.

## 1. Introduction

High-level athletes need to participate in various forms of competitions. After the competitions, systematic physical fitness and special skill training are also required, and daily training is often at a higher intensity level [1–3]. This kind of long-term, high-intensity training will cause the physiological load of athletes. Always in a state of high load, it is easy to cause sports injuries to athletes. Conversely, athletes will not be able to effectively complete high-intensity training due to sports injuries. They will not be able to achieve good results in the competition, which will affect the improvement of athletes' performance. From a physical fitness point of view, high-level athletes' physical fitness level determines their motor skill level [4]. Therefore, sports skills must be based on physical fitness, and sprint sports have very high requirements for athletes' physical fitness. The sprint is one of the highest-intensity sports, and ATP-CP mainly

completes its energy supply form. Sprint athletes often carry out high-intensity sports training, which inevitably brings about a high risk of sports injuries [5]. Studies have shown that the main sports injuries that sprinters often experience are iliac crest bundle friction syndrome, ligament injury, knee injury, ankle injury, hip injury, waist injury, tenosynovitis, abrasions, contusions, and so forth. Suppose that these sports injuries cannot be effectively avoided. In that case, they will continue to affect the systematic high-intensity training of sprinters and shorten the sports life of outstanding athletes. Therefore, athletes' sports injury has become the main factor that continuously limits the performance improvement of high-level athletes. It requires the use of necessary sports injury detection methods to detect high-level athletes [6–8].

The key to sports injury rehabilitation intervention is injury prediction, and the key to prediction lies in accurate recognition of human movements [9]. The action recognition

method based on human posture is the most intuitive method, which uses the law of the change of human posture between consecutive frames to realize human action recognition. This method first recognizes the human body in the video image to obtain the human skeleton information and then uses the human skeleton information for action recognition. Among them, the accuracy of human body gesture recognition directly affects human body action recognition accuracy. Therefore, human body gesture recognition has become an important task of computer vision. Human body gesture recognition aims to identify the human body's joint points in the image and generate the human skeleton. At present, there are two main types of human body recognition methods:

- (i) The top-down method that first recognizes the number of human bodies in the image and then recognizes the human body joint points
- (ii) The bottom-up method that first recognizes the human body joint points in the image and then distinguishes the node attribution

With the development of the Internet of Things and 5G networks, in the case of the Internet of Everything, multi-sensor data fusion is bound to be an important direction for sports injury recognition and prediction in complex environments [10]. Although there have been much researches on sensor-based athlete action recognition, there are still many challenges. To optimize the structure of the recognition model, under the condition of limited parameter scale, the model's feature extraction ability is further improved, thereby improving the recognition of the system's accuracy and practicality. Therefore, this paper constructs a sports injury recognition algorithm based on visual analysis and neural network technology [11–15], which provides auxiliary conditions for rehabilitation intervention.

The main contributions of this paper are as follows:

- (1) This paper proposes a novel sports injury prediction algorithm based on visual analysis technology, which uses multisource sensors to obtain athlete's action data in a complex environment. It uses deep neural networks to predict injury.
- (2) We combine the structure and advantages of CNN networks, such as residual network and hole convolution, and propose a new multihole convolution kernel residual module, which improves the model's ability to extract features at different scales while effectively controlling the parameter scale.
- (3) We performed automatic fusion of multisource sensor data and proposed an automatic multisource sensor data fusion network based on a single-sensor-based behavior recognition algorithm. It automatically performs a feature-level fusion of each sensor's data. Simulation experiments were conducted on the set to evaluate the performance of the proposed method.
- (4) We extensively performed the experiments and the results show that the proposed data fusion network is effective.

The remainder of the paper is organized as follows. In Section 2, related work is studied, followed by the proposed methodology in Section 3. The experimental setup is given in Section 4, and Section 5 concludes the paper.

## 2. Related Work

Competitive sports are characterized by high antagonism and high intensity. The main causes of sports injuries for competitive athletes are overtraining, unreasonable competition training arrangements, physical fatigue, and other factors. According to relevant studies, the injury rate caused by training peaked at 53.4%, chronic injury caused by excessive special training accounted for 23.5%, and sports injury caused by violations of training science accounted for 16.7%.

From the research on the types of sports injuries, chronic injuries, acute injuries, and closed injuries are more common in sports. Acute injuries are manifested as contusions, strains, sprains, fractures, dislocations, and other types of injuries, especially muscle strains. Joint sprains are the most common. From the point of view of sports injuries' characteristics in sprints, the main manifestations are thigh muscle strain, rotator cuff injury, knee sprain, ankle sprain, wrist injury, calf gastrocnemius strain, and ankle tenosynovitis. The most common injuries are back thighs muscle strains and wrist injuries. High-intensity resistance training and sprint events require accelerated movement in the shortest time. This explosive force is the main reason for these two types of injuries [16].

The recognition models used in sports injury recognition systems can be roughly divided into two categories. One is based on classic machine learning algorithms, and the other uses deep learning algorithms such as convolutional neural networks [17–21]. Commonly used classic machine learning algorithms include decision trees, K-nearest neighbors, hidden Markov models, and support vector machines. Scholars at home and abroad have used these algorithms to do a lot of research on sensor-based motion damage recognition. For example, Lee et al. [22] collected angular velocity data through a gyroscope attached to the foot. They used a decision tree model to classify behaviors such as walking, running, going upstairs, and going downstairs. Ignatov et al. [23] proposed an online time-series segmentation method and realized the classification of six behaviors using principal component analysis and the KNN algorithm. Based on the KNN algorithm, Preece et al. [24] compared the performances of 14 feature extraction methods. These features are mainly obtained from the time domain, frequency domain, and wavelet transform. Fleury et al. [25] used accelerometers, magnetometers, and infrared sensors to collect data, performed principal component analysis on the extracted features to obtain ten main features, and realized 35 behaviors' recognition training called the multi-SVM model. The accuracy rate reaches 86%. However, these classic algorithms all require complex and time-consuming feature engineering, manually designing the extracted features, and performing feature selection or dimensionality reduction to filter out representative features.

Computer vision methods based on deep learning avoid this step of feature engineering.

### 3. Methodology

In this section, the details about the methodology are given.

**3.1. Sensor-Based Motion Recognition.** This section provides detailed discussion about the sensor-based motion recognition and its impact on the proposed algorithm.

**3.1.1. Mathematical Description.** Sensor-based behavior recognition can be regarded as a typical pattern recognition problem. Through calculation methods, samples are classified into certain categories according to their characteristics. It is described in mathematical language as follows: Suppose that the user performs certain behaviors belonging to a set  $A$ , and the time-series matrix collected by multiple sensors is  $S$ ; then

$$\begin{aligned} A &= \{A_1, A_2, \dots, A_m\} \\ S &= (d_1, d_2, \dots, d_t, \dots, d_n), \end{aligned} \quad (1)$$

where  $m$  represents the number of action types,  $d_t$  represents the column vector composed of the data collected by the sensor at a time  $t$ , and  $n$  represents the length of the sequence. The time series  $S$  is input into model  $F$ , and, after calculation, the predicted action category sequence  $A$  is obtained, and the real action category sequence is  $\hat{A}$ .

$$\begin{aligned} \hat{A} &= \{\hat{A}_j\}_{j=1}^n = F(S), \quad \hat{A}_j \in A, \\ A^* &= \{A_j^*\}_{j=1}^n, \quad A_j^* \in A. \end{aligned} \quad (2)$$

The action recognition system aims to reduce the difference between the predicted category  $A$  and the real category  $\hat{A}$  by training the model. The purpose is to characterize the difference between the two categories by establishing a loss function  $L$ , and the problem is converted to minimizing the loss function through training.

Usually, the original time series  $S$  is not directly input into the training model but goes through data preprocessing feature extraction and feature selection. The function represents these processes  $\Phi$ : the mapping transformation from the original data  $d_t \in S$  to the feature vector  $\Phi(d_t) \in R^d$ . The final goal of model training is to minimize the loss function  $L(F(\Phi(d_t)), A^*)$ . The symbol denotes that when using sensor-based athlete action recognition, many important objects must be considered, as shown in Table 1.

**3.1.2. Time-Series Imaging Algorithm.** Most of the raw data collected by the sensor is a one-dimensional time series. Applying a two-dimensional convolutional neural network usually needs to be converted into a data format similar to a two-dimensional image. The Gramian Angular Field algorithm is introduced below. The Gramian Angular Field algorithm can convert a one-dimensional time series into a two-dimensional image. The specific implementation steps

are as follows. Suppose that the time series collected by a certain sensor is  $X = \{x_1, x_2, \dots, x_i, \dots, x_N\}$ , which contains  $N$  observations. First, we use normalization operations to make all the values  $X$  between  $[-1,1]$  and  $[0,1]$ . The calculation equation is as follows:

$$\begin{aligned} \tilde{x}_{-1}^i &= \frac{(x_i - \max(X)) + (x_i - \min(X))}{\max(X) - \min(X)}, \\ \tilde{x}_0^i &= \frac{x_i - \min(X)}{\max(X) - \min(X)}. \end{aligned} \quad (3)$$

Then, transform the time series  $X$  from the Cartesian coordinate system to the polar coordinate system.

$$\begin{cases} \varphi_i = \arccos(\tilde{x}_i), & -1 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X}, \\ r_i = \frac{i}{N}, & i \in N. \end{cases} \quad (4)$$

The arc cosine of the normalized observation value  $\tilde{x}_i$  is taken as the angle  $\varphi$  in the polar coordinate system. The time label  $(i/N)$  is taken as the radius in the polar coordinate system. After the two normalization operations, the data have different angle ranges after being converted to the polar coordinate system. The angle of the cosine function corresponding to the data in the range of  $[0,1]$  is  $[0, (\pi/2)]$ , and the angle corresponding to the data in the range of  $[-1,1]$  is  $[0, \pi]$ .

As time increases, the sequence value changes from the original amplitude change to the angle change in the polar coordinate system. By calculating the sum/difference of the trigonometric functions between the points, the time correlation between the points can be identified from the angle of view. Define Gramian Angular Summation Field (GASF) and Gramian Angular Difference Field (GADF) as follows:

$$\begin{aligned} \text{GASF} &= \begin{bmatrix} \cos(\varphi_1 + \varphi_1) & \cdots & \cos(\varphi_1 + \varphi_n) \\ \cos(\varphi_2 + \varphi_1) & \cdots & \cos(\varphi_2 + \varphi_n) \\ \cdots & \ddots & \cdots \\ \cos(\varphi_n + \varphi_1) & \cdots & \cos(\varphi_n + \varphi_n) \end{bmatrix}, \\ \text{GADF} &= \begin{bmatrix} \sin(\varphi_1 - \varphi_1) & \cdots & \sin(\varphi_1 - \varphi_n) \\ \sin(\varphi_2 - \varphi_1) & \cdots & \sin(\varphi_2 - \varphi_n) \\ \cdots & \ddots & \cdots \\ \sin(\varphi_n - \varphi_1) & \cdots & \sin(\varphi_n - \varphi_n) \end{bmatrix}. \end{aligned} \quad (5)$$

### 3.2. Multidilated Convolution Kernel Residual Module

**3.2.1. ResNet.** The core idea of ResNet is to alleviate the difficulty of training caused by the deep network by adding shortcut connections to the network. Before the emergence of ResNet, networks trained too deeply usually had degradation problem and gradients vanishing/exploding. The structure of the residual module is shown in Figure 1.

Suppose that, before adding the shortcut connection, the mapping of the residual module's parameter layer to input  $x$  is  $F_1(x)$ , and the output is  $H_1(x)$ . After adding the shortcut

TABLE 1: Some important symbols in the mathematical description of athlete action recognition.

Symbol	Description
$d_t$	Raw data collected by the sensor
$\Phi$	Preprocessing of raw data and feature extraction and selection
$F$	The recognition model used to predict recognition results
$L$	The loss function determines the optimization direction of the model

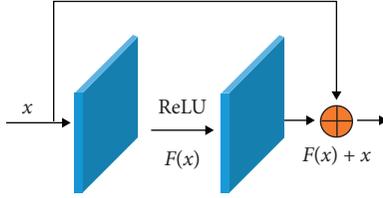


FIGURE 1: Schematic of the residual module structure.

connection, the mapping of the parameter layer of the residual module to input  $x$  is  $F_2(x)$ , and the output is  $H_2(x)$ ; then the outputs are

$$\begin{aligned} H_1(x) &= F_1(x), \\ H_2(x) &= F_2(x) + x. \end{aligned} \quad (6)$$

When the neural network has reached the best performance, the added residual module should learn identity mapping to not reduce the performance of the deep network. It is equivalent to making  $H_1(x) = F_1(x) = x$  and making  $H_2(x) = F_2(x) + x = x$  much simpler because the former needs to realize  $F_1(x) = x$ , while the latter only needs to realize  $F_2(x) = 0$ . The former requires a lot of adjustment and learning of parameters, and a little change in the parameters may cause huge changes to the entire network. The latter only needs to set the parameters to 0 to achieve the performance of the original network. A little parameter change on this basis may improve the performance of the network. In general, the added residual module is easy to learn the identity mapping, so at least it will not worsen the performance of the deep network. At the same time, more parameters make the model's fitting potential greater.

**3.2.2. Dilated Convolution.** The traditional convolutional neural network comprises convolutional layers and pooling layers continuously stacked to deepen the number of network layers and improve performance. However, with the continuous deepening of the network, the pooling layer's existence makes the picture size smaller and smaller. When the picture needs to be further processed subsequently, such as semantic segmentation, it is necessary to restore the picture to the original size through upsampling. In this process, some information will inevitably be lost. Since image segmentation is a pixel-level operation, the result will be affected, and the accuracy will be reduced. By inserting holes in the original convolution kernel, it is possible to increase the receptive field without increasing parameters without pooling and improving network performance.

The principle of hole convolution is shown in Figure 2(a) which is a standard convolution kernel; Figures 2(b) and 2(c)

are, respectively, convolution kernel inserting a certain number of holes. Judging from the input image's direction, the so-called hole is to use the sampling frequency as the parameter rate on the image. For sampling, there is no change when rate = 0; that is, no information is lost, as shown in Figure 2(a). When rate > 1, sample every (rate-1) point on the original image. Figure 2(b) is the case where rate = 2. The red points can be regarded as sampling points on the original image after sampling and convolution. From the convolution kernel's perspective, the hole convolution is to insert (rate-1) zeros in the adjacent points of the traditional convolution kernel and then convolve with the original image. It can be seen that after the 3\*3 convolution kernel is inserted into the hole of rate = 1, its receptive field is equivalent to a 7\*7 convolution kernel, and the number of parameters has not increased at the same time.

**3.2.3. Our Mdc-Res.** In view of the structure and advantages of the several above-mentioned networks, based on the research of residual module and hole convolution, this paper proposes a new Mdc-Res module, the structure of which is shown in Figure 3. The Mdc-Res module mainly includes 4 parts, namely,  $1 \times 1$  convolution, stacked 2 normal  $3 \times 3$  convolutions (i.e.,  $3 \times 3$  convolution with an expansion rate of 1), stacked 2  $3 \times 3$  convolutions with an expansion rate of 2, and stacked 2  $3 \times 3$  convolutions with an expansion rate of 4. The network's upper layer feature map is input to the Mdc-Res module, and the calculations of the above parts are, respectively, processed. Then the 4 calculation results and the input feature map are added to form a quick connection, and, finally, the output of the Mdc-Res module is obtained. In the calculation process, the convolution hyperparameters of the 4 parts are all set to "padding = same, strides = 1" to ensure that the dimensions of the output and input are consistent.

In the Mdc-Res module, by using multiple normal convolution kernels and hollow convolution kernels at the same time, the width of the network is increased, and the ability of the network to extract features of different scales is improved. Simultaneously, since the  $5 \times 5$  convolution kernel is not included, it effectively regulates the model's parameters cut down on computation time. In addition, the shortcut connection in residual learning is introduced into the Mdc-Res module, which suppresses the common degradation problems and gradient disappearance phenomenon in deep neural networks, increases the depth of the network, and improves the fitting ability of the model.

**3.3. Multisensor Motion Recognition.** Multisource sensor fusion can be divided into competition and cooperation.

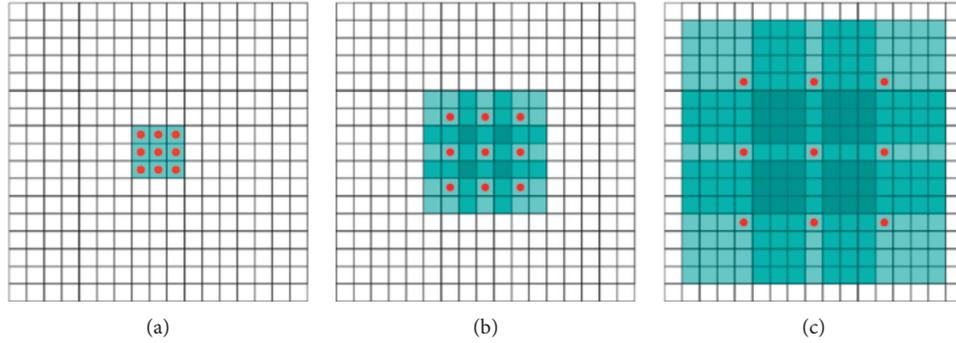


FIGURE 2: Schematic diagram of dilated convolution.

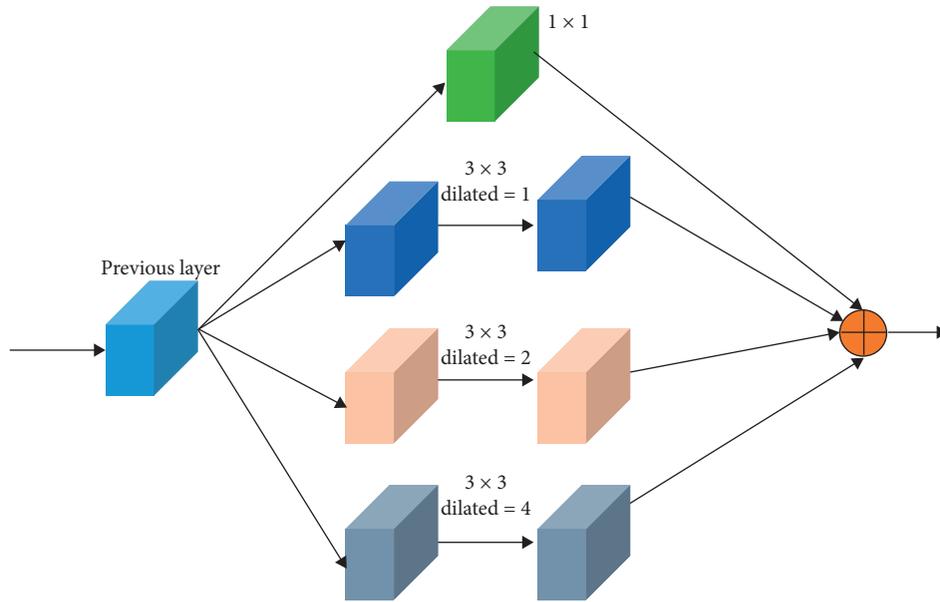


FIGURE 3: Schematic diagram of our Mdc-Res.

Competitive fusion means that multiple equivalent sensors will be used to obtain redundant information and self-calibration. Competitive fusion is relatively rare even if multiple sensors of the same type are used; they are usually placed in different locations on the athlete, so each sensor can provide supplementary information. Cooperative fusion means that each sensor can collect different aspects of the observed object and complement each other to obtain more comprehensive information, thereby improving the system’s accuracy and reliability.

The structure of multi-Mdc-Res is shown in Figure 4, which is mainly composed of  $1 \times 1$  convolutional layer, Mdk-Res module, maximum pooling layer, global average pooling layer, and SoftMax layer. The main difference between it and Mdc-ResNet is that it has multiple data sources, while Mdc-ResNet only processes data from one sensor. The data of multiple data sources is input to multi-Mdc-Res. It is processed by multiple stacked  $1 \times 1$  convolutional layers, Mdc-Res modules, and maximum pooling layer (optional) to automatically extract each

data source’s data. The feature fusion layer then processes the feature to form a new feature map. The new feature map is again processed by multiple  $1 \times 1$  convolutional layers, Mdc-Res module, and maximum pooling layer (optional) to extract effective and robust features from the fused features. Finally, the recognition result is obtained through the calculation of the GAP layer and the SoftMax layer. Except for the activation function of the SoftMax layer that uses the SoftMax function, the other convolutional layers all use the ReLU activation function.

## 4. Experiments

*4.1. Experimental Environment.* The system’s hardware environment in the experiment is CPU Intel Core i7-4700MQ, 2.4GHZ, 16 GB of memory, and the development platform is Python3.6 with Windows 7 operating system. Additionally, Table 2 shows the specific parameters of the experimental environment.

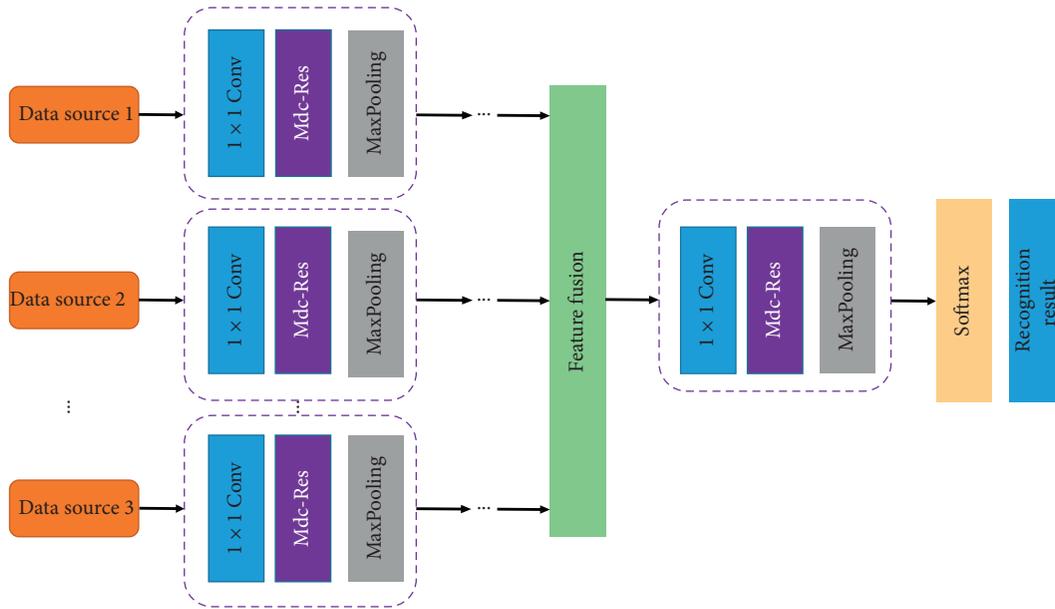


FIGURE 4: Schematic diagram of our multi-Mdc-Res.

TABLE 2: Experimental hardware platform and software simulation environment.

CPU	Intel Core i7-4700MQ @ 2.40 GHz
RAM	16.00 GB
Operating system	Windows 7
Development environment	Python3.6

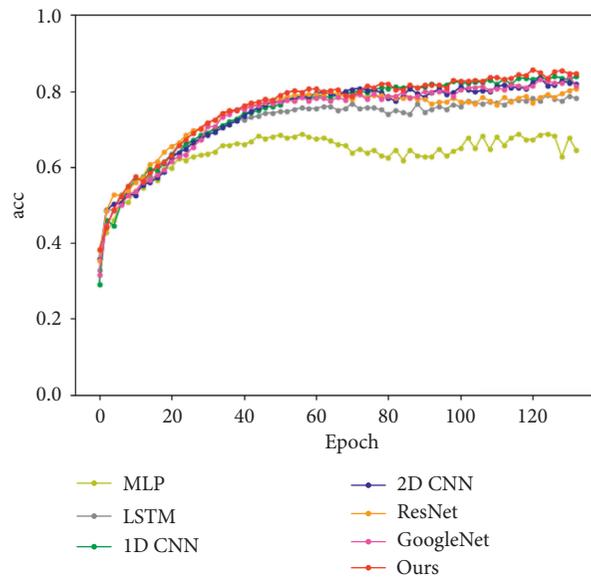


FIGURE 5: The ACC change curve of each model's training set on the UCI HAR data set.

**4.2. Data Set.** The Nonlinear Complex Systems Laboratory of the University of Genoa, Italy, released the UCI HAR data set. The experiment consisted of 30 volunteers between the ages of 19 and 48. They fixed their smartphones to their waists and used the embedded accelerometers and gyroscopes in the phones to collect 6 types of behavioral data (walking, upstairs, downstairs, sitting, standing, and lying). The sampling rate is 50 Hz. This data set has undergone sliding window processing, with a window length of 128, 50% overlap, a total of 10299 samples, and random division, of which 70% is used as the training set and 30% is the test set.

**4.3. Evaluation Methods.** Sensor-based athlete action recognition can be regarded as a multiclassification problem. To evaluate such algorithms' performances, commonly used indicators include Accuracy, Precision, Recall, and F-Measure. Take the binary classification problem as an example to define each indicator. According to the true category of the sample and the model prediction result, the sample is divided into four categories: TP, FP, FN, and TN. The calculation equation is as follows:

$$\begin{aligned}
 \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\
 \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\
 \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\
 F &= \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{(\beta^2 \times \text{Precision}) + \text{Recall}},
 \end{aligned}
 \tag{7}$$

where TP represents the sample itself as the foreground class, and the prediction is correct. TN represents the sample itself as the foreground class, but the prediction is wrong. FP represents the sample itself as the background class, but the prediction is the foreground class. FN represents the sample itself as the background class, and, at the same time, the prediction is correct.

**4.4. Experimental Results.** We have conducted experiments on 7 types of networks, MLP, LSTM, 1D CNN, 2D CNN, ResNet, GoogLeNet, and Fusion-Mdk-ResNet, on the UCI HAR data set. Figures 5 and 6 are the change curves of the model trained for 100 iterations under each evaluation index. It can be seen from the figures above that multi-Mdc-Res has the highest accuracy on the training set, the lowest loss function, and the best model effect. GoogLeNet, ResNet, and 2D CNN are the next best ones. The difference between the three is small, 1D. The effects of CNN, LSTM, and MLP are ranked last. However, in the verification set, the difference between GoogLeNet, ResNet, and Fusion-Mdk-ResNet is not big, which may be due to the small number of verification sets or the deviation of the data distribution of the verification set and the training set, and Figure 7 shows the confusion matrix.

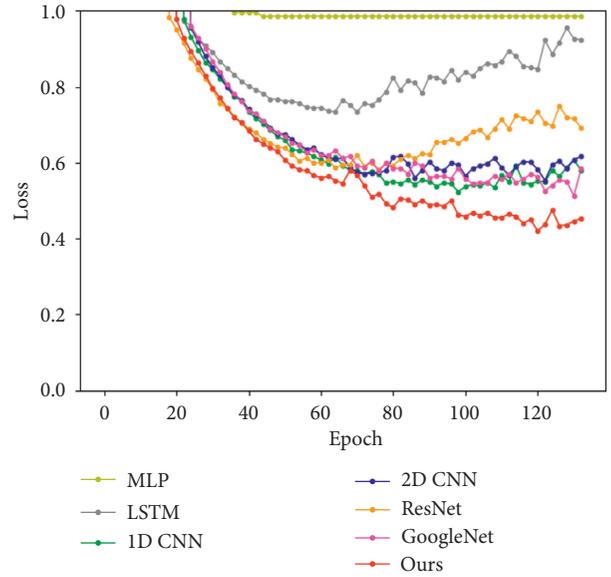


FIGURE 6: The LOSS change curve of each model's training set on the UCI HAR data set.

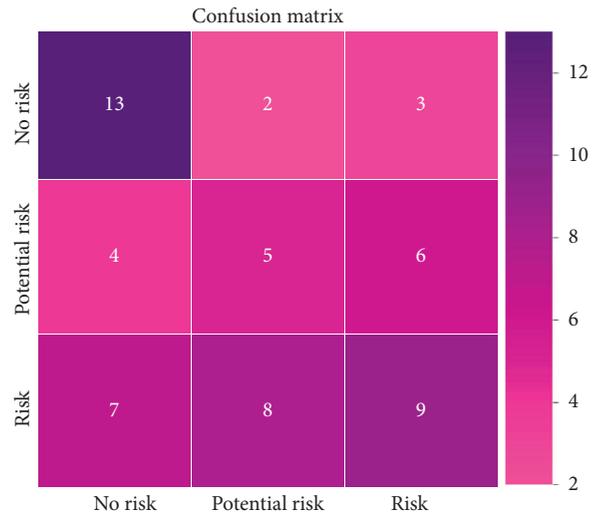


FIGURE 7: Confusion matrix.

## 5. Conclusion

In this paper, we propose a sports injury prediction algorithm based on visual analysis technology. First, the time-frequency graph of sensor data is used as the CNN model's input to study the two-dimensional imaging of one-dimensional time series and convert the one-dimensional time series collected by the sensor into two-dimensional images. Use Gram angle domain algorithm. Observing sensor data provides a new perspective and provides a basis for better use of convolutional neural networks and computer vision technology. Second, combining the structure and advantages of CNN networks such as residual network and hole convolution, a new residual module of porous convolution kernel is proposed, which improves the model's ability to extract features at different scales while effectively

controlling parameters scale. At the same time, based on this module, a sensor-based motion recognition algorithm for athletes is proposed. Many comparative experiments were carried out on the public data set containing only the acceleration sensor to verify the proposed algorithm's effectiveness.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] N. Walter, L. Nikoleizig, and D. Alfermann, "Effects of self-talk training on competitive anxiety, self-efficacy, volitional skills, and performance: an intervention study with junior sub-elite athletes," *Sports*, vol. 7, no. 6, p. 148, 2019.
- [2] M. Gross, Z. E. Moore, F. L. Gardner, A. T. Wolanin, R. Pess, and D. R. Marks, "An empirical examination comparing the mindfulness-acceptance-commitment approach and psychological skills training for the mental health and sport performance of female student athletes," *International Journal of Sport and Exercise Psychology*, vol. 16, no. 4, pp. 431–451, 2018.
- [3] A. Mirifar, J. Beckmann, and F. Ehrlenspiel, "Neurofeedback as supplementary training for optimizing athletes' performance: a systematic review with implications for future research," *Neuroscience & Biobehavioral Reviews*, vol. 75, pp. 419–432, 2017.
- [4] R. A. B. Tedesqui and B. W. Young, "Associations between self-control, practice, and skill level in sport expertise development," *Research Quarterly for Exercise and Sport*, vol. 88, no. 1, pp. 108–113, 2017.
- [5] E. Franchini, S. Cormack, and M. Y. Takito, "Effects of high-intensity interval training on olympic combat sports athletes' performance and physiological adaptation: a systematic review," *Journal of Strength and Conditioning Research*, vol. 33, no. 1, pp. 242–252, 2019.
- [6] C. Huang and L. Jiang, "Data monitoring and sports injury prediction model based on embedded system and machine learning algorithm," *Microprocessors and Microsystems*, vol. 81, Article ID 103654, 2021.
- [7] I. Fernández-Cuevas, J. Arnáiz Lastras, V. Escamilla Galindo, and P. Gómez Carmona, "Infrared thermography for the detection of injury in sports medicine," *Biological and Medical Physics, Biomedical Engineering*, vol. 29, pp. 81–109, 2017.
- [8] S. Jauhiainen, J.-P. Kauppi, M. Leppänen et al., "New machine learning approach for detection of injury risk factors in young team sport athletes," *International Journal of Sports Medicine*, vol. 42, no. 02, pp. 175–182, 2021.
- [9] A. N. Pavlov, D. S. Grishina, A. E. Runnova et al., "Recognition of electroencephalographic patterns related to human movements or mental intentions with multiresolution analysis," *Chaos, Solitons & Fractals*, vol. 126, pp. 230–235, 2019.
- [10] G. Elumalai and R. Ramakrishnan, "A novel approach to monitor and maintain database about physiological parameters of (Javelin) athletes using Internet of Things (IoT)," *Wireless Personal Communications*, vol. 111, no. 1, pp. 343–355, 2020.
- [11] 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.
- [12] W. Cai and Z. Wei, "PiiGAN: generative adversarial networks for pluralistic image inpainting," *IEEE Access*, vol. 8, pp. 48451–48463, 2020.
- [13] X. Ning, S. Xu, W. Li, and S. Nie, "FEGAN: flexible and efficient face editing with pre-trained generator," *IEEE Access*, vol. 8, pp. 65340–65350, 2020.
- [14] 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.
- [15] X. Ning, F. Nan, S. Xu, L. Yu, and L. Zhang, "Multi-view frontal face image generation A survey," *Concurrency and Computation: Practice and Experience*, Article ID e6147, 2020.
- [16] A. Arnold, C. A. Thigpen, P. F. Beattie, M. J. Kissenberth, and E. Shanley, "Overuse physicial injuries in youth athletes," *Sports Health: A Multidisciplinary Approach*, vol. 9, no. 2, pp. 139–147, 2017.
- [17] 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*, vol. 80, no. 7, pp. 11291–11312, 2021.
- [18] X. Ning, W. Li, M. Wei, L. Sun, and X. Dong, "Face anti-spoofing based on deep stack generalization networks, 2018.
- [19] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.
- [20] 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.
- [21] X. Ning, X. Wang, S Xu et al., "A Review of Research on Co-Training," *Concurrency and Computation: Practice and Experience*, 2021.
- [22] M. W. McCarthy, D. A. James, J. B. Lee, and D. D. Rowlands, "Decision-tree-based human activity classification algorithm using single-channel foot-mounted gyroscope," *Electronics Letters*, vol. 51, no. 9, pp. 675–676, 2015.
- [23] A. D. Ignatov and V. V. Strijov, "Human activity recognition using quasiperiodic time series collected from a single tri-axial accelerometer," *Multimedia Tools and Applications*, vol. 75, no. 12, pp. 7257–7270, 2016.
- [24] S. J. Preece, J. Y. Goulermas, L. P. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 871–879, 2008.
- [25] A. Fleury, M. Vacher, and N. Noury, "SVM-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 274–283, 2009.