Recognition of Running Gait of Track and Field Athletes Based on Convolutional Neural Network

Qiusheng Lin[1](✉) and Jin Wang[2]

[1] Guangzhou Huali College, Guangzhou 511325, China
linqiusheng9985@126.com
[2] State Grid Shanghai Municipal Electric Power Company, Shanghai 200120, China

Abstract. With the continuous development of competitive sports, higher requirements have been put forward for the athletic level and technical movements of track and field athletes. Running gait is a key factor that affects the competitive level and technical action of athletes. Therefore, a research on the recognition method of running gait of track and field athletes based on convolutional neural network is proposed. According to the internal noise type of running gait image, the least mean square filtering algorithm is selected as the preprocessing method of running gait image, and it is used to remove the noise of running gait image. Based on this, the LB P feature, Hu moment invariant feature and Haar like feature are extracted as the running gait characteristics, and the convolutional neural network model is designed. With the above designed convolutional neural network model as a tool, the running gait recognition program of track and field athletes is formulated, And determine the calculation formula of the relevant parameters, in order to obtain accurate results of track and field athletes' running gait recognition. The experimental data show that the maximum accuracy of running gait recognition of track and field athletes obtained by using the proposed method is 94%, which fully proves that the proposed method has better performance in running gait recognition.

Keywords: Gait Recognition · Athletes · Convolution Neural Network · Running Gait · Feature Extraction · Recognition Accuracy

1 Introduction

Athletics refers to the general term of all round sports consisting of walking, running, jumping, throwing and other sports as well as other partial events. Track and field sports have a long history, which originated from the basic survival and life activities of human beings. The earliest track and field competition was held in the ancient Greek village of Olympia in 776 BC. Since then, track and field has become one of the official events [1]. By 648, the Olympic Games had added jumping, bidding guns, discus throwing and other events. In 1894, a modern Olympic Games organization was established in Paris, France.

B. Wang et al. (Eds.): ICMTEL 2023, LNICST 533, pp. 224–238, 2024.
https://doi.org/10.1007/978-3-031-50574-4_16

In 1896, the first modern Olympic Games was held in Athens, Greece. Twelve events, including walking, running, jumping and throwing, were listed as the main events of the conference. The successful holding of the 1st Olympic Games marked the establishment of a modern track and field sports system.

Humans have acquired the ability to walk upright through the alternate support of their feet. As the most stressed tissue organ, the pressure on their feet during the whole process of landing, supporting and leaving the ground transits from heel to forefoot, and finally to big toe. According to some research data, the supporting reaction force of an adult's foot during walking can reach 1.5 times of his weight, and it can reach 3–4 times of his weight during running. Therefore, whether the force on the foot is appropriate for the rationality of its activities plays an important role. Generally speaking, most of the current research focuses on the analysis of foot movement through gait analysis. The plantar pressure distribution is a reflection of the whole body posture control and foot structure function. As a mainstream foot research method, plantar pressure analysis has been applied in many fields, such as sports biomechanics, rehabilitation medicine, shoemaking and orthopaedic surgery. The plantar pressure analysis can not only enable us to have a deep understanding of the causes of normal walking gait, but also be an important reference in competitive sports, national fitness and other fields, and can also provide a normal baseline for the research of foot pressure of sick feet.

Athletes are prone to problems in their plantar pressure performance, such as pain caused by excessive local force, potential sports injury risk caused by excessive foot turnover, etc., because they have been trained with large amount of exercise and high intensity for a long time. At present, most of the research directions at home and abroad are aimed at athletic competition and injury rehabilitation, and few of them analyze the gait characteristics and plantar pressure of track and field athletes at ordinary times. In the domestic literature on gait characteristics and plantar pressure of track and field athletes, their experiments often have certain deficiencies for various reasons, such as the selected research objects cannot accurately represent the research group or only a small sample of a few individuals, the indicators measured by the equipment are not consistent with the international research standards, and there is no in-depth study on the impact of plantar pressure and gait characteristics on sports injuries [2]. Therefore, this study takes track and field athletes as the research object. Through the measurement and analysis of the plantar pressure distribution and gait characteristics, it can more accurately and more carefully evaluate the movement mode of the feet and even lower limbs of track and field athletes, find out the relationship between the pressure, gait and sports injury, and the risk factors of lower limb sports injury, and accordingly propose personalized correction plans, Guide track and field athletes to form correct gait and provide adequate protection for athletes' feet.

In the grand occasion of the development of competitive sports, the competitive level and technical movements of track and field athletes around the world are improving and perfecting day by day. At present, it is believed that in order to achieve good sports results, it is necessary to emphasize the economy of running, increase the stride frequency while maintaining a larger stride length, and maintain a stable body weight during running. Therefore, how to improve the scientificity and rationality of athletes' running gait is the key to enhance athletes' competitive level. The existing methods can not accurately identify the running gait of athletes due to the backward application means, so the research on the recognition method of running gait of track and field athletes based on convolutional neural network is proposed.

2 Research on the Recognition Method of Running Gait of Track and Field Athletes

2.1 Running Gait Image Preprocessing

The images of running gait of track and field athletes are obtained through cameras or mobile phones equipped with CMOS devices. Due to the existence of shooting conditions, device defects, recording equipment, image content and other factors, it is inevitable that the image data implementing the recognition method will have different levels of noise interference. The direct visual effects caused by image noise interference usually include: image structure information destruction, boundary pollution, detail loss, image pixel value distribution complexity deepening, etc. These influences will directly cause great trouble to the subsequent running gait feature extraction, increase the complexity of the operation, and then lead to a sharp decline in the recognition accuracy. Choosing an appropriate image preprocessing method for running gait recognition can help improve the data quality, reduce the processing difficulty of subsequent research, and greatly improve the efficiency of running gait recognition.

In the running gait image, the image noise mainly comes from the image acquisition or transmission process, and is affected by the performance of the imaging sensor, the environmental conditions in the image acquisition process, the quality of the sensor components and other factors. For example, when using a charge coupled device (CCD) camera to capture a constructed image, the light level and sensor temperature are the main factors that cause the amount of noise in the resulting image. Image data is polluted in transmission, usually due to the interference in the transmission channel, and then noise is introduced to pollute the image. For example, image transmission under the condition of using radio network transmission is vulnerable to noise interference caused by light or other weather factors.

There are generally three types of noise classification methods, including additive noise and multiplicative noise, external noise and internal noise, stationary noise and internal noise. In the process of image denoising, noise is mainly distinguished by its corresponding noise characteristics, which are usually divided into two categories: spatial characteristics and frequency characteristics. The parameters that define the spatial characteristics of noise and whether the noise is related to the image are discussed. The image frequency characteristic refers to the frequency content of the noise in the Fourier domain (that is, the frequency relative to the electromagnetic spectrum) [3]. In addition to the spatial periodic noise, the academia usually assumes that the noise is independent of the spatial coordinates, and the noise is not related to the image itself (that is, the pixel value is not related to the value of the noise component). Generally, the common noise models in running gait images are Gaussian noise, Poisson noise, salt and pepper noise, etc.

Among them, Gaussian noise is the noise category whose image noise distribution conforms to the Gaussian distribution. Gaussian distribution was first used to describe the probability distribution of random fluctuations in physical processes, named after Karl Friedrich Gauss, a German physicist. Gaussian distribution in image noise is mainly used to describe the distribution of resistance performance change process caused by voltage change. When the image acquisition device undergoes electrical variation and conforms to the Gaussian distribution, it is considered that the image may be polluted by Gaussian noise. The probability density function of Gaussian random variable Z is given by the following formula:

$$P(Z) = \left(\frac{1}{\sqrt{2\sigma}}\right)^{-(Z-\overline{Z})/2\sigma^2} \tag{1}$$

In formula (1), $P(Z)$ represents the probability density function expression of Gaussian random variable Z; Z represents the gray value of running gait image; σ represents the standard deviation of the gray value of the running gait image; $\overline{Z}$ represents the average gray value of the running gait image.

Poisson noise is also a kind of noise interference caused by electronic devices, which can be simulated by Poisson model. In electronics, shot noise originates from the discrete nature of charges. The shot noise also exists in the photon counting process of optical devices. The shot noise is related to the particle properties of light. The reason for the existence of shot noise is that light and current are composed of discrete particle beams [4]. The size of shot noise increases according to the square root of the expected number of events, such as the intensity of current or light. However, as the strength of the signal itself increases faster, the relative proportion of shot noise decreases and the signal-to-noise ratio increases. Therefore, shot noise is most common at low current or low light intensity. Poisson noise is mainly controlled by Poisson distribution model, and the specific formula is as follows:

$$F(\alpha; \lambda) = \frac{\lambda^{\alpha} e^{-\lambda}}{\alpha} \tag{2}$$

In formula (2), $F(\alpha; \lambda)$ represents the Poisson noise distribution model expression; α represents the auxiliary parameters; λ represents the average number of random events per unit time; e represents the Poisson noise distribution coefficient.

According to the type of internal noise in the running gait image, the minimum mean square filtering algorithm is selected to remove the noise, improve the signal to noise ratio of the running gait image, and provide some convenience for the subsequent running gait feature extraction. The minimum mean square error filtering algorithm is based on the fact that both image and noise are random variables. The goal is to find an estimate $\hat{f}(x, y)$ of polluted image $f(x, y)$, and find the optimal filtering template by minimizing the mean square error between them. The error measure formula is:

$$E^2 = H\left\{\left[f(x, y) - \hat{f}(x, y)\right]^2\right\} \tag{3}$$

In formula (3), E^2 represents the error; $H\{\cdot\}$ represents the parameter expected value.

Based on the calculation results of formula (3), determine the optimal filter template, and the expression is:

$$\zeta(x, y) = \varepsilon^0 \left[\frac{1}{\beta(x, y)} \cdot \frac{|\beta(x, y)|^2}{|\beta(x, y)|^2 + \frac{\chi(x,y)}{\delta(x,y)[P(Z)+F(\alpha;\lambda)]}} \right] \tag{4}$$

In formula (4), $\zeta(x, y)$ represents the optimal filter template; ε^0 represents the scale factor with the value range [0, 1]; $\beta(x, y)$ represents the degradation function; $\chi(x, y)$ represents the power spectrum of noise; $\delta(x, y)$ represents the power spectrum of the undegraded image.

The optimal filter template $\zeta(x, y)$ determined by formula (4) is applied to remove the noise of running gait image, the expression:

$$g(x, y) = \frac{f(x, y)}{9 * \zeta(x, y)} + E^2 \tag{5}$$

In formula (5), $g(x, y)$ represents a running gait image after noise removal.

The above process completes the pre-processing of the running gait image, that is, removes the noise in the running gait image, improves the overall signal-to-noise ratio and clarity of the image, and lays a solid foundation for the subsequent running gait feature extraction.

2.2 Running Gait Feature Extraction

Based on the above preprocessed running gait image $g(x, y)$, the relevant parameters of running gait are defined, and the running gait features are extracted to provide a basis for the subsequent design of convolutional neural network model.

In the teaching of track and field sports, the running stage is generally divided into supporting stage and flying stage. The supporting phase includes: landing phase (the moment when the swinging leg touches the ground) to landing phase (the moment when the supporting leg leaves the ground); The flying phase includes: the phase of the support leg leaving the ground to the phase of landing. The parameters involved in the running gait of track and field athletes are as follows:

First, time parameters:

A gait cycle is just the time from foot following to heel landing on the same leg. The normal gait cycle can be divided into two phases: the supporting phase and the swinging phase.

Support phase: from heel landing to toe off the ground, the contact time between the foot and the support surface accounts for about 60% of the gait cycle.
Swing phase: from toe to heel landing, the time of foot leaving the supporting surface accounts for about 40% of the gait cycle.

Second, spatial parameters:

The stride feature of gait is the space feature of foot landing, including step length, stride length, stride width and stride angle.

Step length: linear distance from one foot to the other [5]. Step length is related to height. The higher the height, the greater the step length.
Step length: the linear distance between the same leg and the heel again. The stride length of a normal person is twice the stride length, about 100–160 cm.
Step width: the width between the two foot travel lines.
Step angle: the included angle between the line from the middle of the heel to the second toe and the line of travel, generally less than dexterity.
Step frequency: refers to the number of steps taken per minute when walking.

Third, plantar pressure zone:
The plantar partition is shown in Fig. 1.

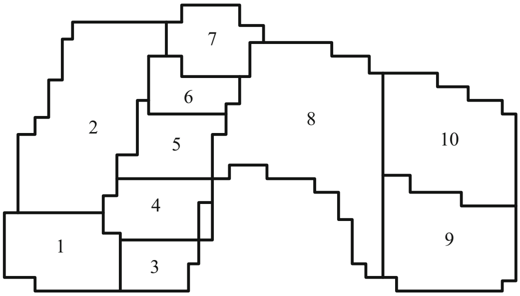

Fig. 1. Schematic diagram of the plantar partition

As shown in Fig. 1, 1 represents the first toe, 2 represents the second to fifth toes, 3 represents the first phalanx, 4 represents the second phalanx, 5 represents the third phalanx, 6 represents the fourth phalanx, 7 represents the fifth phalanx, 8 represents the midfoot, 9 represents the inner heel, and 10 represents the outer heel.

It can be seen from the above analysis that there are many parameters related to running gait. If the running gait feature is extracted on this basis, the amount of calculation is too large to be calculated. Therefore, based on the running gait image $g(x, y)$, LBP feature, Hu moment invariant feature and Haar like feature are extracted as the running gait feature. Among them, the LBP feature is the full name of the local binary feature of the image, which is used to describe the local texture of the image. The original LBP operator is defined through the 3 * 3 matrix window. The threshold value is the template center pixel value. When the pixel value around the center pixel is larger, the surrounding point is marked as 1, otherwise it is 0. After 8 comparisons of 3 * 3 size templates, 8 LBP comparison points can be generated, and then an 8-bit binary number is formed by encoding. The binary number is converted to decimal to obtain the LBP code, that is, the LBP value of the center pixel of the window [6]. This value can be used to further describe the regional texture information for image recognition. The formula is:

$$LBP(x_i, y_i) = \sum_{p=1} 2^p \operatorname{sgn}\left(R_p - R_i\right) \tag{6}$$

In formula (6), $LBP(x_i, y_i)$ represents the LBP value corresponding to the central pixel (x_i, y_i); p represents a random constant. It should be noted that the value is an integer; $\mathrm{sgn}(\cdot)$ represents a symbolic function; R_p and R_i represent the pixel values corresponding to pixels $\left(x_p, y_p\right)$ and (x_i, y_i).

The calculation rules for the symbol function $\mathrm{sgn}(\cdot)$ are as follows:

$$\mathrm{sgn}\left(R_p - R_i\right) = \begin{cases} 1 \; R_p - R_i \geq 0 \\ 0 \; R_p - R_i < 0 \end{cases} \tag{7}$$

The Hu invariant moment is a feature characterizing the shape of an image profile. The definition formula of the $a + b$-order moment of the 2D stochastic strain volume is as follows:

$$G_{ab} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^a y^b g(x, y) dxdy \tag{8}$$

In formula (8), $g(x, y)$ is a running gait image function and a piecewise bounded function. When $g(x, y)$ changes with translation, rotation or scale, the $a + b$ order matrix produces adjoint changes. a and b represents the feature order of the image. To obtain the invariance features, the central moment is defined as follows:

$$\eta_{ab} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(x - \hat{x}\right)^a \left(y - \hat{y}\right)^b g(x, y) dxdy \tag{9}$$

In formula (9), η_{ab} represents the central moment of the running gait image; $\left(\hat{x}, \hat{y}\right)$ represents the central point of the running gait image.

Image scale invariance is obtained by normalization, and on this basis, multiple Hu invariant moments combined are generated, with invariance for scale, translation, and rotation.

Haar like feature was first proposed by Papageorigiou et al., which mainly reflects the change of image gray level, and is a feature description operator commonly used for face description in the field of computer vision research. Haar like features can be calculated from templates of different sizes and composition modes, mainly including linear features, edge features, point features (center features) and diagonal features [7]. Generally, Haar like feature is calculated by integral graph, and the formula is as follows:

$$\sum D = R(1) + R(2) - (R(3) + R(4)) \tag{10}$$

In formula (10), D represents an image block for Haar-like features with four vertices 1, 2, 3 and 4, and in $R(j)$ stores the sum of all pixels in the upper right corner of $R(j)$, and j values 1 to 4.

The above process completes the extraction of running gait features, mainly including LBP features of running gait image, Hu invariant moment features and Haar-like features, providing data support for the identification of the final running gait.

2.3 Convolutional Neural Network Model Design

Based on the above extracted running gait features, a convolutional neural network model is designed to provide a tool support for the subsequent realization of running gait recognition.

The design of convolutional neural network model includes two parts: building CNN layer by layer and training CNN network. The construction process of the network is as follows: The first layer is the input of the network. The CNN network can independently learn the characteristics of the two-dimensional image. The original image is a depth image, which can be directly used as the input of the network. The convolution layer is convolved by $n \times n$ size filters and offset, and characteristic graphs are obtained. All the down sampling layers are obtained by the following processing in turn: summing four pixels per neighborhood, weighting (convolution kernel element), adding offset, and then generating a feature map approximately four times smaller through a sigmoid activation function.

According to the identification of running gait requirements, the convolutional neural network model is designed, with the number of network layers of 5, the number of filters is 3 and 6, respectively, and the filter size is 55 = 25. Specifically, the construction of each layer of networks is explained, as shown in Fig. 2.

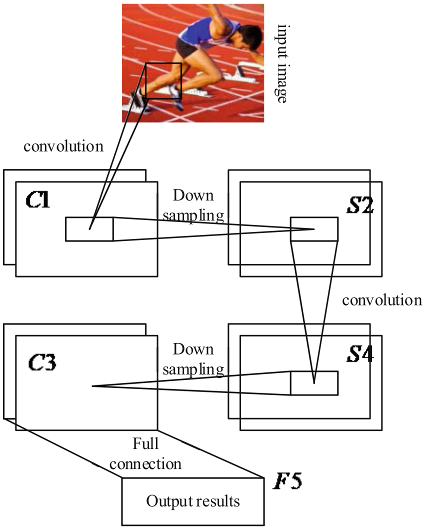

Fig. 2. Process diagram of convolutional neural network model design

As shown in Fig. 2, $C1$ is a convolutional layer, and the first layer mainly extracts the low-level features of the image. Three 5×5-size filters were convolved with the original image to obtain three feature maps, each with a size of 116×92. $C1$ has 7 $\times$ 8 trainable parameters (each filter each has $5 \times 5 = 25$unit parameter and one bias parameter, a total of 3 filters, a total of $(5 \times 5 + 1)\ 3 = 78$ trainable parameters).

$S2$ is a down sampling layer. Down sampling reduces the amount of information processing while retaining useful information according to the principle of image local correlation. Downsampling mainly includes the following methods: average value, maximum value or some linear combination. In this paper, the method of linear combination is used to add four pixels in the neighborhood, multiply them by a trainable parameter, and then add a trainable offset. Therefore, the size of the three feature maps on layer $S2$ is 58×46, $3 \times (1 + 1) = 6$ training parameters and $12 \times 27 \times 21 = 6804$ connections. The lower sampling layer also needs an activation function. The introduction of the activation function is to improve the nonlinear characteristics of the network and decision function. In layer $S2$, sigmoid activation function is selected.

$C3$ is also a convolutional layer, convolved with six 5×5 filters layer $S2$ to obtain six feature maps, each of size 54×42. Similarly, for the calculation of training parameters in layer $C1$, layer $C3$ has $(5 \times 5 + 1)$ 6 trainable parameters.

$S4$ is a subsampling layer. The subsampling was performed in the same way as the $S2$ layer, thus resulting in 27×21 size six feature maps. There are N/m training parameters and $12 \times 27 \times 21 = 6804$ connections. Also there is an activation function, this layer of activation function still selects the sigmoid activation function.

$F5$ is a fully connected layer, with each output cell and all the input units connecting the [8, 9]. The last layer also requires an activation function, in which the ReLu function is selected as the activation function. The final resulting feature graph is then arranged as a column vector to obtain the final eigenvector.

If the number of network layers is more than 5, the following network layers only use convolution layer and the last full connection layer. This is because a large part of useless information has been removed from the previous layers. If the down sampling process is used again, too much information will be lost. If the network structure is more than 5 layers, the number of filters will increase by an equal number sequence with a common ratio of 2.

The above process completes the design of the convolutional neural network model, and introduces its composition in detail, providing a tool support for the follow-up research.

2.4 Development of Running Gait Recognition Procedure for Track and Field Athletes

Using the convolutional neural network model designed above as a tool, the program for recognizing the running gait of track and field athletes is formulated, and the calculation formula of the relevant parameters of the convolutional neural network model is determined, so as to obtain accurate results of recognizing the running gait of track and field athletes.

The whole process of running gait recognition is as follows: collect sample images and mark them, input training samples of each type of running gait into CNN to train the model, adjust the network until convergence: change the output layer into a softmax classifier, input test images, and identify the accuracy of verification results. The running gait recognition program of track and field athletes is shown in Fig. 3.

As shown in Fig. 3, the convolutional neural network model is a key tool for running gait identification in track and field athletes, and its performance is directly related to the accuracy of the identification results. Therefore, the calculation formula for the relevant

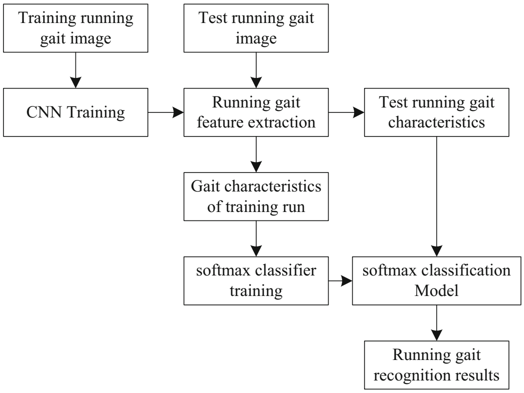

Fig. 3. Running gait recognition program of track and field athletes

parameters of the convolutional neural network model needs to be determined before the program can be executed.

One is the gradient calculation of the convolutional layers.

Assuming that each convolutional layer 1 is connected to a subsampling layer 1 + 1, to obtain the corresponding weight update value for each neuron of layer 1, it is necessary to find the sensitivity ϑ of each neuron node of layer 1 first. For this sensitivity ϑ, it is necessary to sum the sensitivity ϑ^{1+1} of the nodes in the next layer first, and then multiply it by the weight corresponding to these connections. Multiplied by the derivative value of the activation function q' of the input u of the neuron node in the current layer 1, the sensitivity ϑ^{1} corresponding to each neural node in the current layer 1 is obtained.

However, because of the existence of down sampling, the sensitivity ϑ corresponding to one pixel of the sampling layer corresponds to one pixel (sampling window size) of the output image of the convolution layer. Therefore, each node of an image in layer 1 is connected to only one node of the corresponding image in layer 1 + 1. In order to effectively calculate the sensitivity of layer 1, it is necessary to upsample the sensitivity image corresponding to the downsample layer, so that the sensitivity image p size is consistent with the image size of the convolution layer, and then multiply the partial derivative of the activation value of the image of layer 1 by the sensitivity image obtained from the upsampling of layer 1 + 1 element by element.

The weights of the subsampled layer image all take the same value u_j^i, and it is a constant. So only multiplying the results obtained from the previous step by a υ_j^i to complete the calculation of layer 1 sensitivity ϑ. The same computation procedure can be repeated for each feature image j in the convolutional layer. Image matching the corresponding subsampling layer yields:

$$\vartheta_j^i = \upsilon_j^{i+1}\left(q'\left(u_j^{i+1}\right) \cdot up\left(\vartheta_j^{i+1}\right)\right) \tag{11}$$

In formula (11), ϑ_j^i represents a feature image of the subsampling layer; $up(\cdot)$ represents an upsampling operation.

The gradient of each convolutional layer, the sensitivity, is calculated by formula (11).

The second is the selection of activation function.

The activation function is introduced to improve the nonlinear characteristics of the network and decision function without affecting the receptive field of the convolution layer. If there is no such layer, the whole network is linear. Activation function is an important part of neural network. Through function transformation of neural network input, appropriate output can be obtained. Generally, activation function injects nonlinear factors into neural network with poor linear expression ability, so that data can be divided under nonlinear conditions, and data can also be sparsely expressed, so that data can be processed more efficiently [10–12].

According to the requirements of running gait recognition, sigmoid activation function and ReLu function are selected. Through the application of activation function, the training speed of convolution neural network is several times faster. Taking a four layer convolutional neural network as an example, comparing the number of iterations required when the conventional activation function reaches 25% of the training error rate, the six iterations of the network using sigmoid activation function and ReLu function make the training error rate reach 25%.

The third is the setting of softmax classifier.

Softmax regression can be understood as a multi class classifier, which can be used in both supervised and unsupervised machine learning. Softmax is a generalization of the logistic regression model to multiple classification problems, in which the class label can take more than two values.

In general logistic regression, assuming that the set of training samples is $\left\{\left(x^1, y^1\right), \cdots, (x^m, y^m)\right\}$, m is the sample label, and the input feature is x^i. Since logistic regression is specific for dichotomy problems, the class marker $y^i \in \{0, 1\}$. The assumption function is as follows:

$$k(x) = \frac{1}{1 + \exp\left(-\psi^T x\right)} \tag{12}$$

In formula (12), $k(x)$ represents the softmax classification function; ψ represents the classification parameter.

Training the parameter ψ enables to minimize the cost function:

$$J(\psi) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^i \log_2 x^i + \left(1 - y^i\right)\log_2\left(1 - \psi \times x^i\right)\right] \tag{13}$$

In formula (13), $J(\psi)$ represents the cost function of the parameter ψ.

For convenience, the symbol ψ is also used to represent all the parameters. When implementing softmax regression, it is convenient to represent ψ with a matrix of $k \times (n + 1)$, obtained by listing $\psi_1, \psi_2, \cdots, \psi_k$ by rows, as follows:

$$\psi = \begin{bmatrix} -\psi_1^T \\ -\psi_2^T \\ \vdots \\ -\psi_k^T \end{bmatrix} \tag{14}$$

Substitute the convolution layer gradient, activation function and softmax classifier determined above into the established track and field athlete's running gait recognition program, and execute the program to obtain the track and field athlete's running gait recognition results, providing assistance for the improvement of track and field athletes' competitive ability [13–15].

3 Experiment and Result Analysis

3.1 Selection of Experimental Objects

In order to verify the application performance of the proposed method, 100 track and field athletes were selected as experimental objects, and images were collected from the continuous monitoring video of track and field athletes during running. The image extraction frame was 10 s. A total of 5000 images were extracted, of which 3000 images were used as the training set, and 2000 images were used as the experimental set. They were randomly divided into 10 experimental groups, as shown in Table 1.

Table 1. Experimental Group Setting Table

Experimental group number	Average stride/cm	Average step frequency/Step/minute
1	100	180
2	112	175
3	105	164
4	101	150
5	120	145
6	135	168
7	144	120
8	150	114
9	160	152
10	139	150

As shown in Table 1, the average step and average step frequency of the 10 experimental groups are quite different, which meets the requirements of the application performance test of the proposed method.

3.2 Convolution Neural Network Model Training

The proposed method designs a convolutional neural network model, which needs to be trained before the experiment to ensure the accurate operation of the convolutional neural network model, so as to obtain more accurate experimental conclusions.

The training process of convolutional neural network includes four steps and is divided into two stages:

Phase I: Forward propagation process

(1) Take a sample from the sample set and input it into the network;
(2) Calculate the corresponding actual output.

In this stage, the input information is transformed layer by layer and transmitted to the output layer. The calculation process performed by the network is actually the point multiplication of the input and the weight matrix of each layer, plus some deviations, to get the final output result.

Phase II: Back propagation process

(1) Calculate the difference between the actual output and the expected output;
(2) The weight matrix is propagated and adjusted in the direction of minimizing the error.

The training process of the network includes forward propagation and back propagation. Forward propagation is mainly about feature extraction and classification calculation. Back propagation is the back feedback of error and the update calculation of weight. After image input, the neurons on all layers should be initialized first. Convolution and sampling are used to extract and map image features. Multiple convolution and sampling processes can be used here. The multi-level extraction process can extract useful information from the image. After feature extraction is completed, the extracted features will be fed back to the full connection layer. The full connection layer contains multiple hidden layers. The result is fed back to the output layer through the transformation and calculation of data information in the hidden layer. The output layer carries out some calculations and gets the test results. Compare the test results with the expected results, and output the classified results if they are consistent. If the test results do not conform to the expected results, the weights and deviations need to be propagated back. From the output layer to the full connection layer and the convolutional sampling layer, it is passed back once. Until each layer gets its own gradient. Then the weight value is updated to start a new round of training process.

The training process of convolutional neural network model is shown in Fig. 4.

The training of the convolutional neural network model is completed to ensure the reliability of the proposed method.

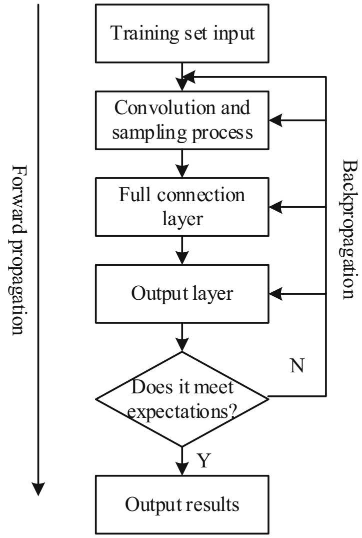

Fig. 4. Flow chart of convolutional neural network model training

4 Analysis of the Experimental Results

Based on the above set experimental conditions and the trained convolutional neural network model, the running gait identification experiment of track and field athletes was conducted, and the running gait recognition accuracy data of track and field athletes are shown in Table 2.

Table 2. Data Table of Track and Field Athletes' Running Gait Recognition Accuracy

Experimental group number	Put forward the method	Minimum limit
1	85%	75%
2	89%	65%
3	90%	60%
4	91%	50%
5	94%	45%
6	89%	62%
7	88%	59%
8	78%	64%
9	80%	70%
10	85%	63%

As shown in the data in Table 2, compared with the given minimum limit value, the recognition accuracy of running gait of track and field athletes obtained by applying the proposed method is higher, with the maximum value of 94%, which fully proves the feasibility and effectiveness of the proposed method.

5 Conclusion

This study introduced convolutional neural network model and proposed a new research on the recognition method of running gait of track and field athletes. The experimental data shows that the proposed method has greatly improved the recognition accuracy of running gait of track and field athletes, which can provide more effective method support for the measurement and improvement of athletic ability of track and field athletes, and also provide some reference and help for related research.

References

1. Qiu, Y., Gao, Z.: Recognition of abnormal gait active image sequences based on low rank decomposition. Comput. Simul. **38**(6), 415–418 (2021)
2. Liu, S., Liu, D., Muhammad, K., Ding, W.: Effective template update mechanism in visual tracking with background clutter. Neurocomputing **458**, 615–625 (2021)
3. Ferguson, E.L.: Multitask convolutional neural network for acoustic localization of a transiting broadband source using a hydrophone array. J. Acoust. Soc. America **150**(1), 248–256 (2021)
4. Dong, S., Jin, Y., Bak, S.J., et al.: Explainable convolutional neural network to investigate the age-related changes in multi-order functional connectivity. Electronics **10**(23), 3020 (2021)
5. Liu, S., Zhu, C.: Jamming recognition based on feature fusion and convolutional neural network. J. Beijing Inst. Technol. **31**(2), 169–177 (2022)
6. Liu, S., Wang, S., Liu, X., et al.: Fuzzy detection aided real-time and robust visual tracking under complex environments. IEEE Trans. Fuzzy Syst. **29**(1), 90–102 (2021)
7. Halász, M., Gerak, J., Bakonyi, P., et al.: Study on the compression effect of clothing on the physiological response of the athlete. Materials **15**(1), 169–169 (2022)
8. Zhu, Z., Yao, C.: Application of attitude tracking algorithm for face recognition based on OpenCV in the intelligent door lock. Comput. Commun. **154**(5), 390–397 (2020)
9. Lai, X., Rau, P.: Has facial recognition technology been misused? A user perception model of facial recognition scenarios. Comput. Hum. Behav. **124**(8), 106894 (2021)
10. Liu, S., et al.: Human memory update strategy: a multi-layer template update mechanism for remote visual monitoring. IEEE Trans. Multimedia **23**, 2188–2198 (2021)
11. Luo, W., Ning, B.: High-dynamic dance motion recognition method based on video visual analysis. Sci. Program. **2022**, 1–9 (2022)
12. Tha, B., Sk, B., Mt, B., et al.: Comparing subject-to-subject transfer learning methods in surface electromyogram-based motion recognition with shallow and deep classifiers. Neurocomputing **489**, 599–612 (2022)
13. Sattar, N.Y., Kausar, Z., Usama, S.A., et al.: FNIRS-based upper limb motion intention recognition using an artificial neural network for transhumeral amputees. Sensors **22**(3), 726–733 (2022)
14. Zhang, K., Zhao, D., Liu, W.: Online vehicle trajectory compression algorithm based on motion pattern recognition. IET Intel. Transp. Syst. **16**(8), 998–1010 (2022)
15. Muhammad, U., Yu, Z., Komulainen, J.: Self-supervised 2D face presentation attack detection via temporal sequence sampling. Pattern Recogn. Lett. **156**(4), 15–22 (2022)