

Evaluation Model of Athletes' Lower Extremity Training Ability Based on LSTM Algorithm

Yuanyuan Gao
Sports Department, Hebei Agricultural
University, China
qianqiang1999@163.com

Qiang Qian
Sports Department, Hebei Agricultural
University, China
nehansom@163.com

Xuefeng Sun
Sports Department, Hebei
Agricultural University, China
sxf198201@163.com

Jiankun An
Baoding Sports School, China
liyuanan@163.com

Yutao Yuan
Sports Department, Hebei Agricultural
University, China
yutaoyuan_yy@163.com

Abstract: To achieve intelligent evaluation of the lower limb movement ability of athletes with sports disabilities, this article selects young athletes and middle-aged and young athletes with sports disabilities as the research objects, and healthy young athletes as the control group. Gait videos, GRF and knee angles of the subjects were collected to extract and analyse gait contours and features. The improved Visual Background extractor (ViBe) algorithm has the highest accuracy of 0.978 in PETS2006 video sequences; the recall rate of three algorithms in Highway video sequence is the highest, and the recall rate of improved ViBe algorithm is the highest, up to 0.965. At $\epsilon=0.7$, the accuracy of the improved myloss training set is higher than that corresponding to other ϵ values; when the number of iterations is 98, the accuracy rate of improved myloss training set is 0.963, while the accuracy rate corresponding to the cross entropy loss function is 0.945. When the number of iterations is 151, the accuracy rate of the Xception LSTM model is 0.956, higher than that of other models. Among corresponding mean \pm standard deviation (GSA-MS) indicators, the GSA-MS values of Group L are significantly higher than those of group N and group Z ($P<0.001$). The correlation between GAS indicators and Gait Abnormality Rating Scale (GARS-M) is strong, with a correlation coefficient of 0.90.

Keywords: Short-term memory network, athletes, lower extremity motor ability, evaluation model.

Received December 13, 2022; accepted June 19, 2023

<https://doi.org/10.34028/iajit/21/1/13>

1. Introduction

In recent years, the country's support for the sports industry has been increasing, and becoming a professional athlete is a choice for many people. In order to bring glory to their country and realize their life value in the competition, they have been conducting sports training for years and months. High intensity training has brought considerable pressure to their bodies, leading to many athletes suffering from occupational diseases and a significant decline in their athletic performance, hindering their development [8, 14]. Therefore, conducting intelligent diagnosis and treatment can provide a preliminary understanding of the physical condition of athletes, reduce the pressure on doctors, and enable athletes to receive timely and effective treatment. And for athletes who engage in lower limb exercises, timely mastering their lower limb exercise abilities can help them adjust their training plans in a timely manner and have a certain promoting effect on their mental recover [16]. In the evaluation process, classification algorithms are involved the research found that combining Long and Short Term Memory (LSTM) with Convolutional Neural Network

(CNN) can help improve the classification accuracy of users' comment emotions, distinguish different comment emotions, and have good classification performance [13]. Face pose is diverse. By combining LSTM and CNN and using incremental clustering strategy, different facial poses can be recognized with good performance [2]. Considering the good performance of LSTM and CNN in dynamic video classification, the article applies this method to the evaluation of lower limb motor ability of athletes with sports disorders, aiming to automate the evaluation of lower limb motor ability of athletes with sports disorders, provide valuable reference data for doctors' diagnosis and treatment, and promote the progress of intelligent diagnosis and treatment.

2. Related Works

Pollen *et al.* [11] tracked and analyzed the functional exercise screening scores of different groups of athletes to evaluate their athletic abilities. Related analysis has found that the level of competition will not have an impact on the comprehensive score of functional sports screening for these athletes. The research review has

certain reference value in predicting athlete injuries. Higgins *et al.* [3] performed shoulder arthroscopy on professional athletes from different organizations, recorded relevant data on athletes recovering from competitions before and after the examination, and scored accordingly based on their performance. Analysis shows that the probability of professional athletes participating in recovery competitions after surgery is moderate. Pandya *et al.* [10] conducted a case study on injuries in young professional athletes, compiled literature on fifth bone fractures, and investigated factors affecting athlete recovery. The case analysis results show that effective splinting fixation for athletes and urging them to engage in early grading activities is beneficial for their recovery. Lebedev *et al.* [5] constructed a health dynamic health system in athlete health monitoring, recorded relevant health data of athletes, and developed electronic passports for each athlete. From relevant application analysis, it can be concluded that the system is feasible. Nunes *et al.* [9] conducted a study on the relevant references for evaluating the health status of adolescent football players and analyzed the influencing factors of their cardiovascular health indicators. After analysis, it is known that weight indicators can have an impact on cardiovascular health indicators. Maria-Sacheli *et al.* [7] analyzed the relationship between pathological motor function limitations and motor imagination skills in athletes with motor disabilities, evaluated their motor performance, and detected relevant neural signals. In the comparative experimental results, dynamic gait imagination is beneficial for the treatment of athletes. Jones *et al.* [4] conducted testing and analysis on the identification of injury risks in adolescent football sports, and found that combination testing methods including functional exercise screening are beneficial for identifying non-contact injury risks.

Ercolano and Rossi [1] combined CNN and LSTM to form a hybrid algorithm to identify real-time daily activities. In the process of limb related feature extraction, the hybrid algorithm has better recognition effect and higher recognition accuracy, which provides reference data for social assisted robot design. Manju *et al.* [6] studied the prediction and recognition of suspicious activities, automatically recognized human activities, and integrated Three-Dimensional Convolutional Neural Network (3DCNN) and LSTM to make them predict future abnormal events by identifying human activities in videos. After detection, the fusion algorithm has high recognition accuracy in the recognition of criminal data sets. In order to realize the diagnosis of plant diseases and pests, Verma and Dubey [17] processed relevant mobile video images by using LSTM and simple recurrent neural network to extract image data features and other related processing. The results show that the method used in the study has good prediction ability. To obtain the correct frame in the surveillance video, Radarapu *et al.* [12] propose to

apply LSTM to extract the text of video features, which shows that attention mechanism is widely used in video subtitle work. Su *et al.* [15] faced with the problem of face tampering in the video, proposed a deep forgery video detection method, and used LSTM to extract the spatial and temporal information of the video. The results show that the detection accuracy of this method is high. Yadav *et al.* [18] used the deep convolution long and short term memory network to detect human activity and human fall behavior. The results show that the method has good detection effect and high accuracy. Zhang *et al.* [19] classified the video time through a two-stage neural network. Under the effect of the maximum mean pool, they converted the frame sequence into anchor points and overcame the gradient disappearance through LSTM. The results show that the classification effect is good.

To sum up, sports disabled athletes are a concern for many scholars, and most of them focus their research on athlete rehabilitation, with relatively little research on the evaluation of athlete lower limb motor ability. In addition, LSTM and CNN have good performance in classification and recognition. Therefore, this article uses LSTM and CNN to explore the evaluation of lower limb motor ability in athletes with sports disabilities.

3. Evaluation Model of Athletes' Lower Limb Training Ability Based on LSTM Algorithm

3.1. Multimodal Data Acquisition and Processing of Human Lower Limbs

Sports injury is a common problem for many athletes, which makes them have sports obstacles. Relevant scholars help them recover their sports ability by using exoskeleton equipment. However, before making the equipment, it is necessary to master the athletes' lower limb sports ability. Therefore, this paper collects and processes the multi-modal data of the lower limbs of athletes with sports disabilities. First of all, the movement state of human lower limbs is analyzed. With the natural standing as a reference, the center of gravity of the human body is taken as the base point, the direction perpendicular to the front of the human body is taken as the sagittal axis, the direction perpendicular to the sagittal axis on the left side of the human body is taken as the coronal axis, and the direction of the top of the head is taken as the vertical axis. Three basic planes are formed, including sagittal plane, coronal plane and horizontal plane. The reference space is determined according to the basic plane and the basic axis. Relevant research points out that when people walk, their lower limb movements mainly focus on the sagittal plane, and lower limb movements are carried out with the cooperation of various parts of the lower limb. Therefore, when collecting the lower limb joint angle, the hip joint, knee joint and ankle joint motion angles in

the sagittal plane are selected, which are θ_h , θ_k , θ_a . According to the human body structure, θ_k is always negative; when the hip joint flexes forward, θ_h is positive, and vice versa; When the ankle joint does dorsiflexion upward, θ_a is positive, and vice versa

In the process of walking, the legs swing periodically and circularly. This paper studies a complete gait cycle, which can be divided into swing phase and support phase. The proportion of the two phases in the whole gait cycle is about 40% and 60% respectively. The latter is the starting point of the gait cycle. One side of the heel from landing to lifting the toe to landing is the support phase, and according to the number of feet landing, it can be divided into single support phase and double support phase. In the double support period, the stability of walking will be enhanced, so during the walking process, athletes with sports disabilities tend to have a higher frequency of double support periods. Then, gait information will be collected, and the subjects will voluntarily participate in the collection. Selecting young and middle-aged athletes with motor disabilities as the experimental group, and selecting healthy young athletes as the control group. The number of people in both groups was 20. The modified Gait Abnormality Rating Scale (GARS-M) was used to evaluate the gait abnormalities of subjects in walking videos. There were 7 items in total, with a total score of 0-21. The higher the score, the more severe the gait abnormalities of the subjects. Set the young healthy control group as Group J, the young athletes with sports disorders as Group Q, and the young athletes with sports disorders as Group Z. See Table 1 for the basic information of the three groups.

Table 1. Basic information of subjects.

Grouping	Number of people	Age	Subjects	GARS-M
Group J	20	23.1 ± 1.8	Healthy young athletes	0
Group Q	10	23.7 ± 1.2	Young athletes with sports disabilities	3-9
Group Z	10	33.5 ± 4.8	Young and middle-aged athletes with sports disorders	8-16

After the statistics of the basic information of the three groups of subjects, their motion data were collected. The article used the three-dimensional gait acquisition system (VICON MX) to collect and process the relevant data, and its schematic diagram is shown in Figure 1.

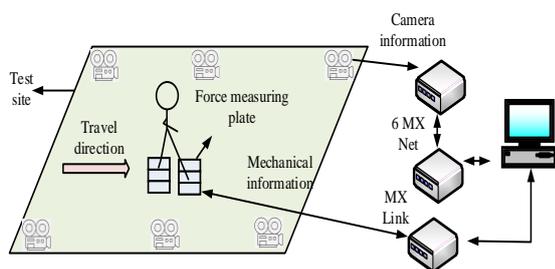


Figure 1. System diagram.

The system has six MX infrared cameras, which can record the movement information and video of the lower limb of the subject in real time. Under the action of the force measuring plate and the reflective marker ball at the lower limb joints, the data is collected and transmitted, and the processed relevant data is obtained after analysis. Before collection, the reflective marker ball needs to be fixed between the lower limb joints and joints of the subject. Then, let it stand in the center of the site for about 5 seconds, input the relevant information of the subject's limbs to the acquisition system, establish a static model, and statically calibrate the output data to ensure accurate data acquisition. The main data collected is the foot pressure during walking, which is also the Ground Reaction Force (GRF). The AMTI OR6-7 force measuring platform is used for synchronous collection. The gait cycle stage can be judged according to the value of GRF. The coordinate system of the force measuring plate is the same as that of the operation information collection system. The lower part of the force measuring plate surface Z_0 is the force measuring origin. The moment at any point on the force measuring plate surface is calculated. The point Q coordinate is set as $(x, y, 0)$. The moment calculation formula at this point is shown in Equation (1).

$$\begin{cases} M_x = F_x * 0 - F_y * Z_0 + F_z * Y + T_x \\ M_y = F_x * Z_0 + F_y * 0 - F_z * X + T_y \\ M_z = -F_x * Y + F_y * X + F_z * 0 + T_z \end{cases} \quad (1)$$

In Equation (1), F represents force; M represents torque; T_x , T_y and T_z represent free torques around X , Y and Z axes respectively; $T_x=0$ and $T_y=0$. According to Equations (1) and (2) can be obtained

$$\begin{cases} M_x = F_z * Y \\ M_y = -F_z * X \\ M_z = -F_x * Y + F_y * X + T_z \end{cases} \quad (2)$$

In Equation (2), F represents force and M represents torque. The collected lower limb motion data is processed and analyzed. When gait video contour is extracted, the Visual Background Extractor (ViBe) algorithm is used. The algorithm is simple, but there are some shortcomings. When foreground objects pass through the ghosted area, the background model updating mechanism will slow down the speed of eliminating the latter, which affects the detection accuracy. Therefore, in this paper, improved ViBe algorithm is used to preprocess gait video to eliminate the influence of environment and clothing. First, carry out background modeling, initialize the background model with a single frame image, assume that adjacent pixels share a similar time distribution, set the pixel value corresponding to position x to $w(x)$, randomly select n pixel values in the neighborhood of $w(x)$, the set of which is $B(x) = \{w_1, w_2, \dots, w_n\}$, then build reference model $B_R(x)$ in the background model to save the initial

frame pixel values and avoid the occurrence of ghost regions. Carry out pixel classification, calculate the distance of each sample value in $w(x)$ and $B(x)$, and determine whether x belongs to the background when the calculated distance is less than the threshold value of R , count the number of corresponding pixel values, which is called the approximate sample size of D . D_{min} is the classification threshold, if $D > D_{min}$, then x is the background point, otherwise it is the front spot. Increase the discrimination conditions to avoid Ghost points. The relevant formula is shown in Equation (3).

$$\begin{cases} w(x, t_0) - w(x, t) > 0 \\ w(x, t_0) > 0 \end{cases} \quad (3)$$

In Equation (3), $w(x, t_0)$ represents the pixel value of x in $B_R(x)$, and $w(x, t)$ represents the pixel value of x in the current image. Then, the background model is randomly updated with each new frame, so that the background model merges the foreground objects that have a long dwell time or stop suddenly. Set the number of times the image pixels are continuously detected as the background to L and when $L > L_{min}$, L_{min} are the threshold values. Modify the previous scenic spots into background points, and update $B_R(x)$. From the obtained kinematics and dynamics data, we can see the changes of different human joint angles and GRF during walking. According to the walking characteristics, representative and differential data is selected as the data needed for the lower limb evaluation model.

3.2. Lower Extremity Motor Function Evaluation Based on Improved LSTM

In image recognition, CNN algorithm has good performance, but it ignores the time dimension information, while LSTM algorithm can avoid the situation of low gradient and explosion when mining time series features. The combination of CNN and LSTM can better learn the spatio-temporal features of video. In addition, Kernel Principal Component Analysis (KPCA) has better performance in nonlinear feature extraction, data dimensionality reduction and other aspects. Therefore, this paper combines Xception LSTM and KPCA for the evaluation of lower limb motor ability. In the LSTM network, assuming that the current time is t , and the input of the previous time is $t-1$, t is x_t , and $t-1$ the output value of the previous time is h_{t-1} . The input gate calculates the input information according to x_t , h_{t-1} , so as to update the model state. At time t , the mathematical expressions of the input gate vector i_t , the forgetting gate vector f_t , and the output gate vector o_t are shown in Equation (4)

$$\begin{cases} i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \end{cases} \quad (4)$$

In Equation (4), b represents offset; b represents coefficient, and the corresponding weight is W . Memory cells integrate the features of time t and memory unit information of time $t-1$, and the relevant mathematical expressions are shown in Equation (5).

$$c_t = i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) + f_t c_{t-1} \quad (5)$$

In Equation (5), c_t represents the state vector at time t and c_{t-1} represents the state vector at time $t-1$. Then, the hidden layer state output h_t is obtained as shown in Equation (6).

$$h_t = o_t \tanh(c_t) \quad (6)$$

In Equation (6), h_t represents the output value at time t . The bidirectional LSTM, which includes output layer, back layer, forward layer, input layer and back layer, is adopted to calculate the hidden layer output from time t to time h_t in reverse. The forward layer is calculated in the opposite way the two directions each get a sequence, and the two groups of columns at each time are combined to get the Equation of 1, as shown in Equation (7).

$$h_t = o_t \tanh(c_t) \quad (7)$$

In Equation (7), \vec{h}_t represents h_t under forward calculation. KPCA can reduce the dimension to process high-dimensional nonlinear data and mine the nonlinear information Assuming that the number of samples in the input space is M , $x_k \in R^N$, $k=1,2,\dots$, use nonlinear mapping ϕ to map the samples and map them to high-order feature space F . The result is $\phi(x_1), \phi(x_2), \dots, \phi(x_M)$. If it is centralized in F , the relevant Covariance matrix can be obtained. In F , the distance between x_{x_μ} and x_{x_ν} is the inner product, resulting in matrix $K_{\mu\nu} = \Phi(x_\mu)\Phi(x_\nu)$ of $M \times M$. On this basis, non-zero eigenvalues λ_k and eigenvectors v^k can be obtained. When the actual application can not meet the centrization conditions, kernel functions can be used to replace the inner product, so that the centralized kernel matrix can be obtained. Its mathematical expression is shown in Equation (8).

$$K_{\mu\nu} \rightarrow K_{\mu\nu} - \frac{1}{M} \left(\sum_{\mu=1}^M K_{k\nu} + \sum_{\nu=1}^M K_{\mu\omega} \right) \frac{1}{M^2} \sum_{k,\omega=1}^M K_{k\omega} \quad (8)$$

In Equation (8), $K_{kv} = \Phi(x_k)\Phi(x_\nu)$, $K_{kv} = \Phi(x_k)\Phi(x_\nu)$, $K_{\mu\omega} = \Phi(x_\mu)\Phi(x_\omega)$. Then, the evaluation model is established. The method is mainly divided into two parts, namely feature extraction and index establishment apply migration learning to network training. Set the pre-training model parameters in ImageNet, a visual database, as the Xception initialization parameters, and then combine them with LSTM to obtain the Xception-LSTM model. Its network structure is shown in Figure 2.

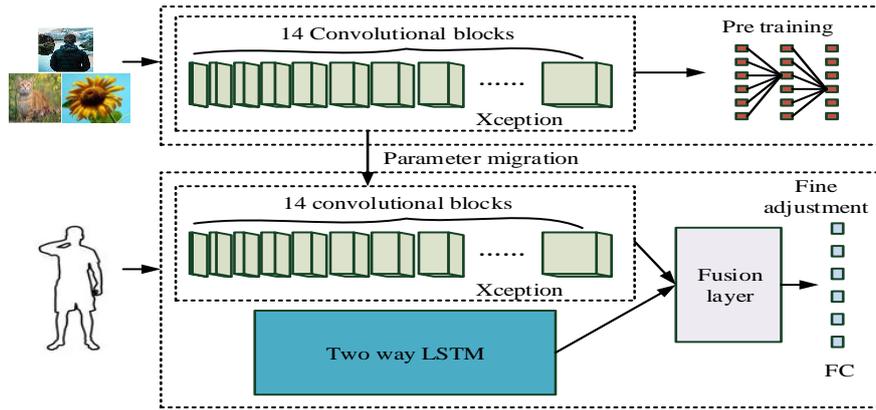


Figure 2. Network structure of the model.

In the model network structure, Xception includes 14 convolutional blocks named Black_1-Black_14. Under the attention mechanism, the fusion layer combines the features of Xception and bidirectional LSTM, and the extracted features are located in the Full Connection

layer (FC). Among them, the model network parameters are shown in Table 2.

Table 2. Model network parameters.

Project	Xception					LSTM
	Black_1	Black_1-Black_4	Black_5-Black_12	Black_13	Black_14	
Composition of network layers	Convolutional layer+RELU Nonlinear activation function+Convolution layer+RELU Nonlinear activation function	Separable convolution	Residual structure	Separable convolution	Depth divisible roll up layer+RELU Nonlinear activation function+Depth divisible convolution layer+RELU Nonlinear activation function+Global average pooling layer	LSTM+LSTM
Input	150*150*3	72*72*64	9*9*728	9*9*728	5*5*1024	150*150*3
Output	72*72*64	9*9*728	9*9*728	5*5*1024	2048	2048
Convolution kernel size	3*3	3*3 1*1	3*3	3*3 1*1	3*3	/
Convolution step	1	1	1	1	1	/
Maximum reddenning	2*2	2*2	2*2	2*2	2*2	/
Minimum pooling	2	2	2	2	2	/

Table 1 contains the parameter information of Xception and bidirectional LSTM. The parameters of the former are initialized through transfer learning, and the convolution block parameters remain unchanged. Only the latter is trained. Then, the model composed of the two is trained, and the network parameters are fine tuned. Under the attention mechanism, the fusion layer combines the features of Xception and bidirectional LSTM. The extracted features are located in FC Feature normalization processes knee joint angle and GRF, and connects the results with features in FC to form fusion features. The related schematic diagram is shown in Figure 3.

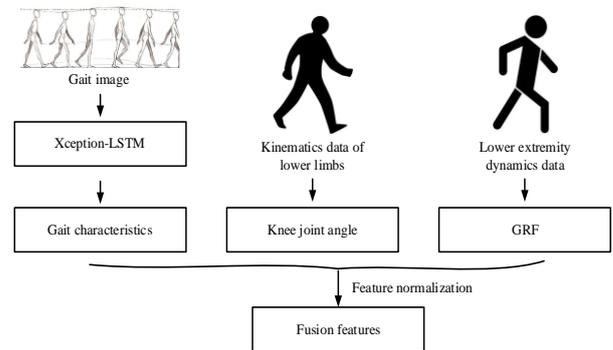


Figure 3. Feature extraction and fusion.

After feature extraction and fusion, the fusion feature matrix $M=[m_1, m_2, \dots, m_k]_{l \times k}$ is normalized. There are k subjects. The dimension after feature fusion is l , and the fusion feature of the i subject is m_i . After normalization, the correlation between indicators of different magnitudes is removed. Calculate the mean value μ_i and variance σ_i of each subject, and set the normalized

M as the fusion matrix M^G , whose range is $[0,1]$. The relevant formula is shown in Equation (9).

$$m_i^G = \frac{1+(m_i - \mu_i)/3\sigma_i}{2} \quad (9)$$

In Equation (9), m_i^G represents normalized m_i . Then, feature dimension reduction and Gait Ability Score (GAS) are established. KPCA is used to remove redundant information of M^G and extract effective features. During this period, Variance Account For (VAF) and threshold E are set. E and VAF are used to determine the feature dimensions after dimension reduction, and the feature dimensions after dimension reduction are set to $[c^1, c^2, \dots, c^k]$. GAR indicators are established through it. The formula involved in the period is shown in Equation (10).

$$VAF = \frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^N \lambda_i} \geq E \quad (10)$$

In Equation (10), λ_i represents the i non-zero eigenvalue, p and N represent the number of non-zero eigenvalues. The subjects' fusion characteristics are c^α , and the average value of N groups' fusion feature vectors is \bar{c}^{TD} , the deviation between c^α and \bar{c}^{TD} is shown in Equation (11).

$$d^{\alpha,TD} = \|c^\alpha - \bar{c}^{TD}\| \quad (11)$$

In Equation (11), \bar{c}^{TD} represents the deviation between \bar{c}^{TD} and \bar{c}^{TD} . Calculate the original GAS index, and its formula is shown in Equation (12).

$$GAS_{raw}^\alpha = \ln(d^{\alpha,TD}) \quad (12)$$

In Equation (12), GAS_{raw}^α represents the original GAS index. Z-Score standardization is carried out for Group N μ_i and σ_i , and the formulas involved are shown in Equation (13).

$$zGAS_{raw}^\alpha = (GAS_{raw}^\alpha - Mean(GAS_{raw}^{TD})) / (S.D.(GAS_{raw}^{TD})) \quad (13)$$

In Equation (13), $zGAS_{raw}^\alpha$ represents the standardization GAS_{raw}^α , and GAS_{raw}^{TD} represents the GAS index of Group N . Quantitative evaluation can be carried out for subjects through Equation (14), as shown in Equation (14).

$$GAS^\alpha = 100 - 10 * zGAS_{raw}^\alpha \quad (14)$$

In Equation (14), when the calculation result is greater than or equal to 100, the lower limb movement ability of the subject is normal, otherwise, the more serious the movement disorder is. In the design of network parameters, the classification of softmax loss function is overconfident, which is easy to lead to over fitting and non-convergence of the network. This function is improved. See Equation (15) for details

$$myloss = -\varepsilon \log\left(\frac{e^{z_i}}{\sum_k e^{z_k}}\right) - (1-\varepsilon) \sum_{i=1}^n \frac{1}{n} \log\left(\frac{e^{z_i}}{\sum_k e^{z_k}}\right) \quad (15)$$

In Equation (15), i represents the category, and its range is $[1,2,\dots,k]$; ε represents the parameter, which is used to balance the two loss functions. Overall, the research method is shown in Figure 4.

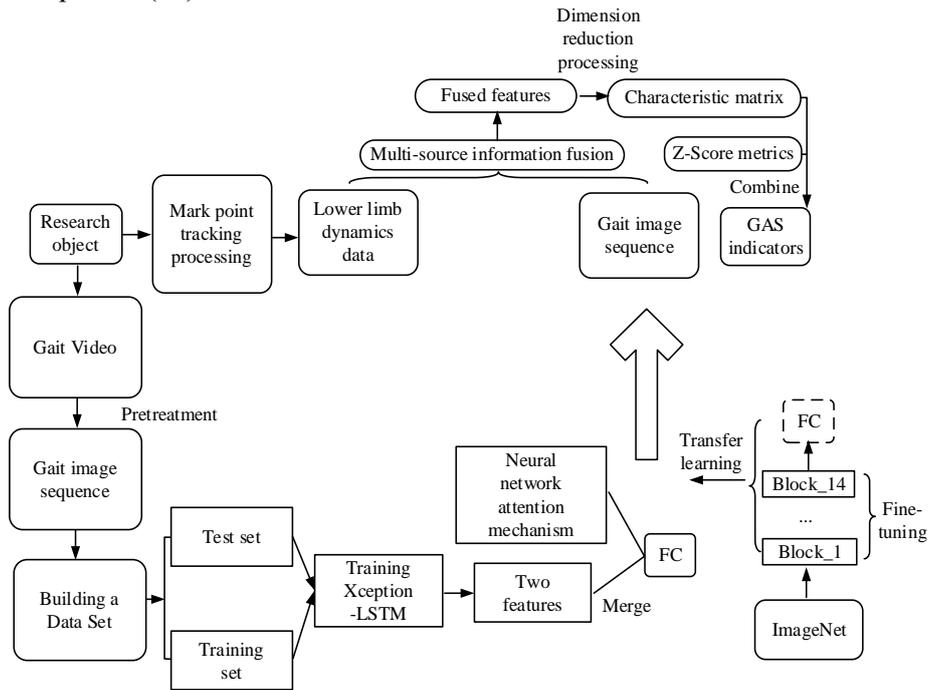


Figure 4. Research plan.

In Figure 4, there are mainly two parts, one is feature extraction. In the process of network training, transfer

learning is introduced, and the model parameters pre trained on ImageNet are taken as the initialization

parameters of Xception. Combine this Xception with LSTM and train the Xception-LSTM model. Use the collected gait image data to fine tune the network parameters. Xception and LSTM generate two features, which are combined with neural network attention mechanisms to obtain FC. In another section, establish indicators. FC, lower limb dynamics and Kinematics data features are normalized and connected in series to form fusion features, and the fusion features are processed by KPCA dimension reduction. Combine the obtained feature matrix with the Z-Score indicator to construct the GAS indicator.

4. Analysis of Experimental Results

To verify the effectiveness of the improved ViBe algorithm in the article, the ViBe algorithm is combined with the three frame difference ViBe algorithm as the comparison algorithm, and the test data set is the Change Detection data set. Four video sequences in the data set are selected as the algorithm test video sequences. When selecting test frames, except that the initial frame needs to contain moving targets, other frames are randomly selected. Select the initial frame of the test sequence Pedestrians as frame 590, with a total of 655 test frames; The initial frame of the Backdoor test sequence is frame 273, with a total of 330 test frames; The initial frame of the test sequence Highway is the 960th frame, with a total of 285 test frames; The initial frame of the test sequence PETS2006 is frame 25, with a total of 720 test frames. Accuracy and recall are used as evaluation indicators. After testing, the recall and accuracy of the three algorithms in the four video sequences are shown in Figure 5.

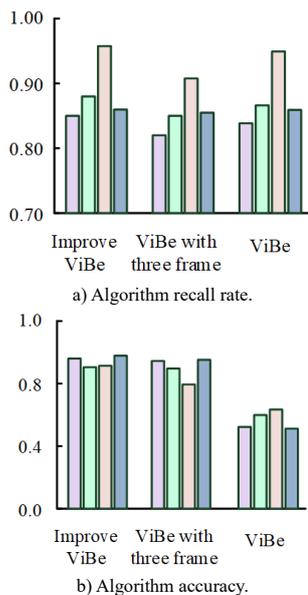


Figure 5. Algorithm recall and accuracy.

In Figure 5-a), the recall rates of the three algorithms in different video sequences are different, and the recall rates in the same video sequence are different. In the PETS2006 video sequence, the recall rate of the

improved ViBe algorithm is 0.864, slightly higher than that of the ViBe algorithm; compared with other video sequences, the recall rate of three algorithms in Highway video sequence is the highest, and the recall rate of the improved ViBe algorithm is the highest, up to 0.965, while the recall rate of the ViBe algorithm is 0.952. In the Pedestrians video sequence, the recall rate of the ViBe algorithm and the improved ViBe algorithm combined with three frame difference is 0.817, 0.835 and 0.836 respectively. In Figure 5b, it is obvious that the precision of the ViBe algorithm is lower than that of the other two algorithms, On the whole, the accuracy of the improved ViBe algorithm is the highest. In the Pedestrians video sequence, the accuracy of the ViBe algorithm is the lowest, reaching 0.523, 0.444 lower than the improved ViBe algorithm. The accuracy of the improved ViBe algorithm is the highest, up to 0.967. Compared with other video sequences, in the PETS2006 video sequence, the accuracy of the improved ViBe algorithm is the highest, reaching 0.978. The lower limb motor function evaluation model is trained and it runs on the deep learning framework keras. The training is divided into two parts. First, only the LSTM network is trained, with the training times of 100 and the batch size of 32. RMSProp is used as the optimization algorithm. Then, thaw and train the network layer's thawed Xception, with the Adam optimization algorithm. The learning rate is 0.0001, and the training times are the same as the training LSTM. The batch size is 16. Process the processed videos of the control group and the experimental group, and use the images obtained as the training set and test set respectively, the number of corresponding pictures is 10324 and 2035, respectively. The data set is expanded through data enhancement to reduce the occurrence of over fitting. Since the improved myloss function will affect the accuracy of network recognition, the experiment shows that the optimal value of ε is 0.7. To verify this result, select $\varepsilon = 0.7$, $\varepsilon = 0.7$ and $\varepsilon = 0.3$, and compare the results of the improved myloss function and cross entropy loss function as shown in Figure 6.

In Figure 6, with different values, the accuracy rate and loss function value of the training set are different. At that time, the accuracy rate of the improved myloss training set was higher than that corresponding to other values, while at that time, the accuracy rate of the improved myloss training set was the lowest; compared with the two loss functions, the improved myloss has a higher correlation accuracy, while the loss function value is just the opposite. With the increase of the number of iterations, the accuracy of the training set increases rapidly first and then tends to be stable. The change trend of the loss function value is opposite to the change trend of the accuracy rate. In Figure 6-a), when the number of iterations is 98, the accuracy rate of the improved myloss training set is 0.963, while the corresponding accuracy rate of the cross entropy loss function is 0.945, 0.018 lower than the former; at this

time, the loss function value of the improved myloss training set is 0.07, 0.11 lower than the cross entropy loss function. Analyze the recognition accuracy of the Xception-LSTM model, select the Visual Geometry Group Network (VGG) and Inception V3 in CNN, combine them with LSTM respectively to form the corresponding recognition model, and use them as the comparison model. Use the KTH dataset to study the accuracy of the model under different iterations, as shown in Figure 7.

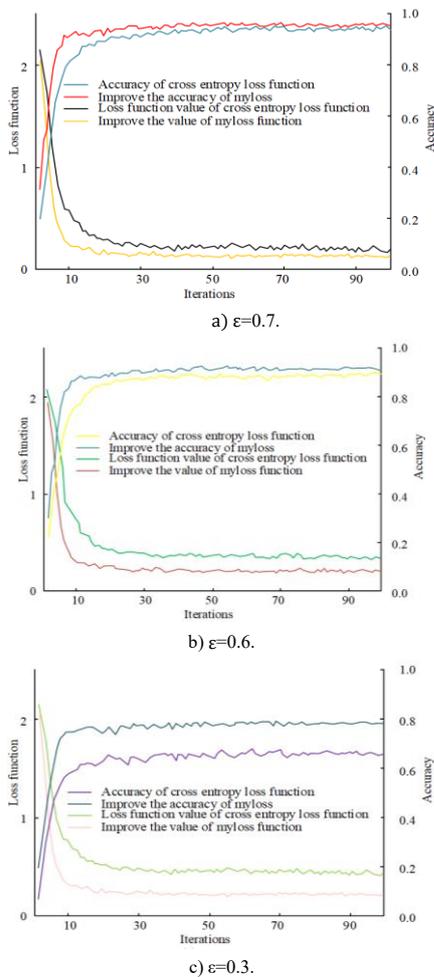


Figure 6. Improved myloss and cross entropy loss function.

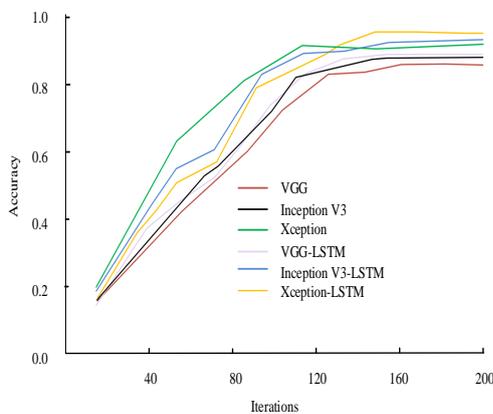


Figure 7. Accuracy of the model.

In Figure 7, as the number of iterations increases, the accuracy of the model increases first and then tends to

be flat; compared with the CNN model, the accuracy rate of the fusion model is higher, while the accuracy rate of the Xception LSTM model is the highest overall. When the number of iterations is 151, the accuracy rate of the Xception LSTM model is 0.956, 0.07 higher than that of the VGG-LSTM model, while the accuracy rate of the Xception model is 0.903. After the evaluation of the Xception LSTM model, the GAS values of different groups of subjects and the corresponding mean \pm standard deviation GSA-MS are shown in Figure 8.

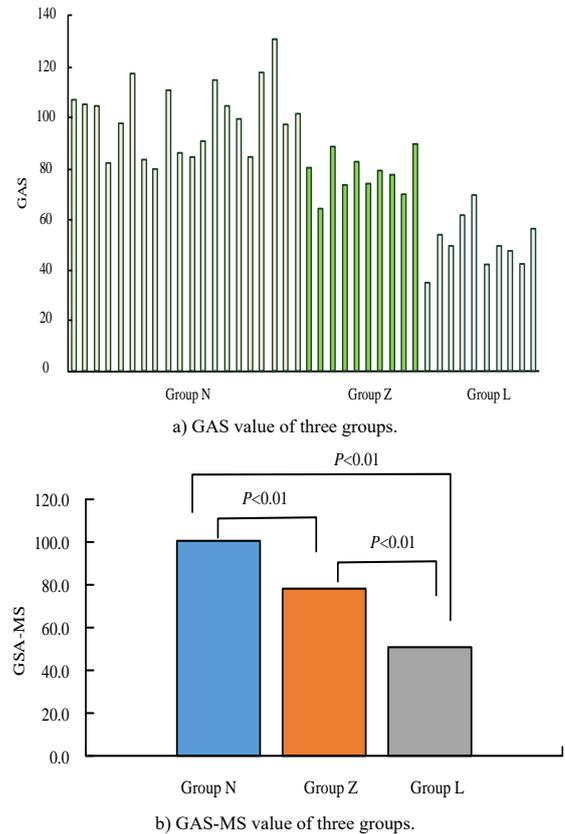


Figure 8. GAS value and GSA-MS value of each group.

In Figure 7, in general, the GAS value of subjects in Group N is higher than that of subjects in Group Z and Group L. The GAS value of subjects in Group Z is the second, while that of subjects in group L is the lowest in Figure 8-a), the highest GAS value of subjects in Group N is 132, and the lowest GAS value is 80, 8 higher than the highest GAS value in Group Z, while the lowest GAS value in group Z is 64; The highest and lowest GAS values of Group L are 69 and 35 respectively. In Figure 8b), the highest mean difference of Group L is 49.753, and the lowest standard deviation is 5.077, followed by group Z; GSA-MS values in group L are significantly higher than those in group N and group Z ($P < 0.001$). To verify the consistency between the GAS indicators established in the article and the evaluation by professional doctors, the GARS-M scale was used to evaluate the movement disorders of group Z and Group L, and the GARS-M values were obtained. The GARS-M values are linearly regressed with the two groups of GAS indicators, and the Principal Component Analysis

(PCA) is used to calculate the GAS. Then, the same linear regression is performed again, and the relevant results are shown in Figure 9.

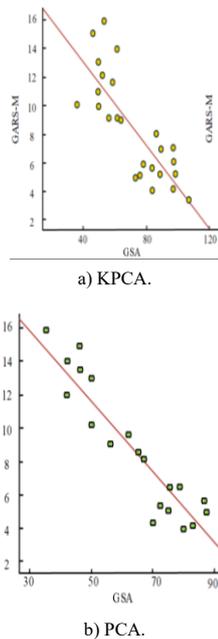


Figure 9. Correlation analysis.

It can be seen from the two sub graphs in Figure 9 that the correlation between GAS indicators and GARS-M is very strong under KPCA method, with a correlation coefficient of 0.90; The correlation coefficient corresponding to PCA method is 0.78, which shows that the dimension reduction effect of KPCA is better than that of PCA.

5. Conclusions

To understand the lower limb movement ability of athletes with sports disabilities and provide effective data for the development of exoskeleton equipment, this article selects healthy young athletes, young athletes with sports disabilities, and middle-aged and young athletes, and divides them into three groups. The subjects' movement data is collected and processed through the VICON MX three-dimensional gait acquisition system, and gait contour images are extracted with the help of the improved ViBe algorithm. After the system data collection, processing and screening, the data required for the evaluation model is obtained. Then, under the action of the Xoption LSTM model combined with KPCA, the GAS value of the subject is finally obtained. Through the verification of relevant results, it is found that compared with other video sequences, the recall rate of three algorithms in Highway video sequences is the highest, and the recall rate of the improved ViBe algorithm is the highest, up to 0.965; in the Pedstrians video sequence, the lowest accuracy of the ViBe algorithm is 0.523, while the highest accuracy of the improved ViBe algorithm is 0.967. The accuracy and loss function values of the

training set are different. When $\varepsilon = 0.7$ is used, the accuracy of the improved myloss training set is higher than that of other ε values. When $\varepsilon = 0.3$ is used, the accuracy of the improved myloss training set is the lowest; compared with the two loss functions, the improved myloss function has a higher correlation accuracy, while the loss function value is just the opposite. Compared with the CNN model, the fusion model has a higher accuracy, and the overall accuracy of the Xoption LSTM model is the highest. In Figure 7-a), the highest GAS value of subjects in group N is 132, and the lowest GAS value is 80, which is higher than the highest GAS value in group Z. 8. The correlation between GAS indicators and GARS-M is strong under the KPCA method, with a correlation coefficient of 0.90. In this paper, a new network structure can be used to improve the calculation speed of the algorithm.

References

- [1] Ercolano G. and Rossi S., "Combining CNN and LSTM for Activity of Daily Living Recognition with a 3D Matrix Skeleton Representation," *Intelligent Service Robotics*, vol. 14, no. 8, pp. 1-11, 2021. DOI:10.1007/s11370-021-00358-7
- [2] Guan Y., Fang J., and Wu X., "Multi-Pose Face Recognition Using Cascade Alignment Network and Incremental Clustering," *Signal, Image and Video Processing*, vol. 15, pp. 63-71, 2021. <https://doi.org/10.1007/s11760-020-01718-z>
- [3] Higgins M., Defroda S., Yang D., Brown S., and Mulcahey M., "Professional Athlete Return to Play and Performance After Shoulder Arthroscopy Varies by Sport," *Arthroscopy Sports Medicine and Rehabilitation*, vol. 2. no. 3, pp. 391-397, 2021. DOI: 10.1016/j.asmr.2020.10.001
- [4] Jones S., Fuller J., Chalmers S., Debenedictis T., and Zacharia A., "Combining Physical performance and Functional Movement Screen Testing to Identify Elite Junior Australian Football Athletes at Risk of Injury," *Scandinavian Journal of Medicine and Science in Sports*, vol. 30 no. 8, pp. 1449-1456, 2020. doi: 10.1111/sms.13686
- [5] Lebedev G., Gureeva A., and Tikhonova Y., "Software System for Dynamic Athlete Health Monitoring," *Procedia Computer Science*, vol. 112 no. 1, pp. 1664-1669, 2017. <https://doi.org/10.1016/j.procs.2017.08.194>
- [6] Manju D., Seetha M., and Sannulal P., "Early Action Prediction using 3DCNN with LSTM and Bidirectional LSTM," *Turkish Journal of Computer and Mathematics Education*, vol. 12 no. 6, pp. 2275-2281, 2021.
- [7] Maria-Sacheli L., Laura Z., Matteo P., Santis C., and Pelosi C., "A Functional Limitation to the Lower Limbs Affects the Neural Bases of Motor Imagery of Gait," *NeuroImage: Clinical*, vol. 20, pp. 177-187, 2018.

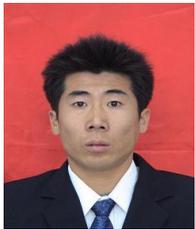
- <https://doi.org/10.1016/j.nicl.2018.07.003>
- [8] Nankervis B., Ferguson L., Gosling C., Storr M., and Ilic D., "How Do Professional Australian Football League (AFL) Players Utilise Social Media During Periods of Injury? A Mixed Methods Analysis," *Journal of Science and Medicine in Sport*, vol. 21 no. 7, pp. 681-685, 2018. DOI: 10.1016/j.jsams.2017.10.034
- [9] Nunes H., Faria E., Martinez P., and Oliveira-Júnior S., "Cardiovascular Health Indicators in Soccer Exercise During Adolescence: Systematic Review," *International Journal of Adolescent Medicine and Health*, vol. 33 no. 3, pp. 53-63, 2021. DOI: 10.1515/ijamh-2020-0301
- [10] Pandya T., Dale S., Donnison E., and Kluzek S., "Conservative Management of Fifth Metacarpal Head Fracture in a Professional Cricketer: A Case Study and Literature Review," *Clinical Case Reports*, vol. 8, no. 9, pp. 1682-1685, 2020. doi: 10.1002/ccr3.2960
- [11] Pollen T., Keitt F., and Trojian T., "Do Normative Composite Scores on the Functional Movement Screen Differ Across High School, Collegiate, and Professional Athletes? A Critical Review," *Clinical Journal of Sport Medicine*, vol. 31, no. 1, pp. 91-102, 2021. DOI: 10.1097/JSM.0000000000000672
- [12] Radarapu R., Gopal A., Madhusudhan N., and Anand K., "Video Summarization and Captioning Using Dynamic Mode Decomposition for Surveillance," *International Journal of Information Technology*, vol. 13, no. 9, pp. 1927-1936, 2021. <https://doi.org/10.1007/s41870-021-00668-0>
- [13] Ramaswamy S., and Chinnappan J., "RecogNet-LSTM+CNN: A Hybrid Network with Attention Mechanism for Aspect Categorization and Sentiment Classification," *Journal of Intelligent Information Systems*, vol. 58, pp. 379-404, 2022. <https://doi.org/10.1007/s10844-021-00692-3>
- [14] Strack S., Macdonald C., Valencia E., and Davison M., "Case for the Specialised Sports Physical Therapist to Be an Essential Part of Professional Athlete Care: Letter from America no. 1," *British Journal of Sports Medicine*, vol. 53, no. 10, pp. 587-588, 2018. DOI: 10.1136/bjsports-2017-097575
- [15] Su Y., Xia H., Liang Q., and Nie W., "Exposing DeepFake Videos Using Attention Based Convolutional LSTM Network," *Neural Processing Letters*, vol. 53, no. 5, pp. 4159-4175, 2021. <https://doi.org/10.1007/s11063-021-10588-6>
- [16] Surbhi K., Akashdeep S., Amandeep V., Vishal D., and Chahat G., "A Comparative Study on Deep Learning and Machine Learning Models for Human Action Recognition in Aerial Videos," *The International Arab Journal of Information Technology*, vol. 20, no. 04, pp. 567 - 574, 2023. <https://doi.org/10.34028/iajit/20/4/2>
- [17] Verma T. and Dubey S., "Prediction of Diseased Rice Plant Using Video Processing and LSTM-Simple Recurrent Neural Network with Comparative Study," *Multimedia Tools and Applications*, vol. 80, no. 19, pp. 29267-29298, 2021. <https://doi.org/10.1007/s11042-021-10889-x>
- [18] Yadav S., Tiwari K., Pandey H., and Akbar S., "Skeleton-based Human Activity Recognition Using ConvLstm and Guided Feature Learning," *Soft Computing: A Fusion of Foundations, Methodologies and Applications*, vol. 26, no. 2, pp. 877-890, 2022. <https://doi.org/10.1007/s00500-021-06238-7>
- [19] Zhang L. and Xiang X., "Video Event Classification Based on Two-Stage Neural Network," *Multimedia Tools and Applications*, vol. 79, no. 4, pp. 1-16, 2020.



Yuanyuan Gao graduated from Xi'an Institute of Physical Education in 2003 with a major in physical education. She obtained a Master's degree in Physical Education and Training from Xi'an Institute of Physical Education in 2006. She is currently a teacher in the Sports Department of Hebei Agricultural University. Her areas of interest are Physical Education Teaching, Training, and Refereeing.



Qiang Qian lecturer, graduated from the National Traditional Sports major of Shenyang Institute of Physical Education in 2003, and has been working in the Sports Department of Agricultural University of Hebei since 2004.



Xuefeng Sun graduated from Tianjin University of Sport in 2004, majoring in sports training. In 2010, he obtained a master's degree in physical education and training from the School of Physical Education of Hebei Normal University. He is currently a teacher in the Sports Department of Hebei Agricultural University. His areas of interest are Physical Education Teaching, Training, and Refereeing.



Jiankun An in 2003 obtained a bachelor's degree in Traditional Chinese Martial Arts from Shenyang Institute of Physical Education. Now he is the martial arts coach of Baoding No.1 Sports School. His fields of interest are Martial Arts, Karate, Wrestling And Sanda.



Yutao Yuan professor, graduated from the Physical Education Department of Hebei Normal University in 1990 and joined the Physical Education Department of Agricultural University of Hebei in the same year. He has won two Hebei Province teaching achievement Awards and three Hebei Province Excellent Social Science Awards.