AMERICAN JOURNAL OF BIOCHEMISTRY AND BIOTECHNOLOGY

RESEARCH ARTICLE

Gait Analysis Combining Sports Biomechanics and Imaging

Chao Zhu

Chongqing Vocational Institute of Safety Technology, Wanzhou 404100, Chongqing, China

*Corresponding Author: 15086726251@163.com

Abstract: Aiming at the problems of marker positioning error, soft tissue interference, and insufficient personalized modeling ability in traditional gait analysis, this study proposes a gait assessment method that integrates sports biomechanics and multimodal imaging analysis. Firstly, based on the public OU-ISIR gait database, a Convolutional Neural Network (CNN) is used to extract the key gait features in the human silhouette map, and a Long Short-Term Memory Network (LSTM) is used to perform temporal modeling of the dynamic gait cycle to reduce the impact of clothing changes and perspective differences on feature extraction. Secondly, a personalized joint dynamics model is constructed through finite element analysis, and an iterative optimization algorithm is used to associate imaging parameters with muscle dynamics indicators, thereby improving the model's ability to adapt to individual anatomical differences. Experimental results show that the classification accuracy of this method under covariate conditions reaches 97.5% (clothing changes) and 91.2% (viewing angle differences). In joint angle prediction, the angle errors of the hip, knee, and ankle joints are 1.1°, 1.2°, and 1.3°, respectively, which are lower than those of the existing methods. Ablation experiments further verify the effectiveness of each module, indicating the key role of dynamic modeling, personalized adaptation, and parameter optimization in improving the overall performance. The results show that the proposed method performs well in terms of accuracy, robustness, and efficiency, providing reliable technical support for the practical application of gait analysis in clinical monitoring and smart devices.

Keywords: Gait Analysis, Sports Biomechanics, Imaging, Convolutional Neural Network, Long Short-Term Memory Network, Iterative Optimization Algorithm

Received: 07-01-2026 | **Revised:** 19-03-2026 | **Accepted:** 15-04-2026 | **DOI:** 10.3844/ajbbsp.2026.22.02.24

Introduction

Gait analysis, as an important research field in sports biomechanics and clinical medicine, is widely used in disease diagnosis, rehabilitation treatment, sports science, intelligent monitoring, and other scenarios. However, traditional gait analysis methods face many challenges in practical applications, such as insufficient accuracy caused by marker positioning errors and soft tissue interference, as well as limited personalized modeling capabilities. These limitations affect the accuracy of gait analysis and also limit its applicability in complex environments. In recent years, with the development of imaging technology and artificial intelligence algorithms, multimodal data fusion has provided a new research direction for gait analysis. By combining sports biomechanics models with imaging data, the errors of traditional methods can be reduced, and accurate adaptation to individual anatomical differences can be achieved, thereby improving the clinical value and universality of gait analysis. Therefore, exploring gait analysis methods that combine sports biomechanics with imaging has important theoretical significance and practical application value.

© 2026 Chao Zhu. This is an open access article published by Science Publications and distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

DOI: 10.3844/ajbbsp.2026.22.02.024

In recent years, gait analysis, as an important research field in sports biomechanics and clinical medicine, has been widely used in disease diagnosis, rehabilitation therapy, sports science, and intelligent monitoring. Feng Yuguo assessed the test-retest reliability of lower limb joint angles and spatiotemporal gait characteristics using the Xsens inertial sensor motion capture system. The findings demonstrated the validity of the Xsens system in assessing the lower limb joint angles and spatiotemporal gait metrics of young, healthy individuals [1]. Tian Haoyu discussed the shortcomings and development direction of current gait research in practical applications and discussed the feasibility of using skeleton data collected by Kinect in gait analysis. Sethi described and comprehensively analyzed the machine learning methods based on marker and marker-free analysis and the insights of clinical gait analysis, providing a basis for future research directions in this field [2]. Stenum used OpenPose to perform gait analysis using videos recorded from multiple different angles, and the results showed a significant improvement in the ability to perform quantitative gait analysis in almost any environment using low-cost equipment and computer vision [3]. Wren used Theia3D to label pediatric clinical gait patients, which had greater trial-to-trial variability [4]. It can be seen that the existing gait analysis methods have obvious deficiencies in marker positioning error, soft tissue interference, and personalized modeling capabilities, resulting in limited accuracy and difficulty in adapting to changing conditions in complex environments.

The CNN algorithm, with its powerful feature extraction capabilities and advantages in capturing multi-scale information, provides new possibilities for solving the above problems, allowing gait analysis to more accurately adapt to the needs of diverse scenarios. Semwal combined LSTM and CNN to train these models on a benchmark public human walking dataset. The proposed LSTM-CNN sequential model was able to generate stable gait trajectories and provide personalized gait reference trajectories for the rehabilitation of amputees and stroke patients [5]. Vidya's experiments using multiple gaits confirmed that the proposed CNN-LSTM classifier model can achieve a maximum multi-class classification accuracy of 98.32% [6]. Berke Erdaş proposed a method based on deep learning, using gait data represented by Quick Response code and a CNN deep learning method for regression, and verified the effectiveness of CNN [7]. Merlin Linda proposed a method based on deep learning, using gait data represented by Quick Response code and a CNN deep learning method for regression, and verified the effectiveness of CNN [8]. The main parameters of gait recognition are speed and clothing changes. Merlin Linda proposed a CNN-based capsule network (CNN-CapsNet) model to solve the differences caused by auxiliary factors and gait recognition. Li designed a multimodal gait abnormality recognition approach using CNN-Bidirectional Long Short-Term Memory (CNN-BiLSTM) network to solve the problems caused by smooth data interference and long time series in gait analysis [9]. With the continuous development of imaging technology and artificial intelligence algorithm, multimodal data fusion opens up a new research direction for gait analysis. If sports biomechanical models are used in combination with imaging data, the error of traditional methods can be reduced, and accurate individual anatomical differences can be achieved, which has important clinical value and universality.

This study aims to propose a gait assessment method integrating sports biomechanics and multimodal imaging analysis, so as to solve the problems of marker positioning error, soft tissue interference, and insufficient personalized modeling ability in traditional gait analysis, and improve the accuracy, robustness, and applicability of gait analysis, providing reliable technical support for its practical application in clinical monitoring and smart devices. Therefore, exploring gait analysis methods that combine sports biomechanics with imaging has important theoretical significance and practical application value.

Materials and Methods

Data Preprocessing

Data preprocessing is a crucial step in gait analysis, and its purpose is to eliminate covariate interference, such as clothing changes and perspective differences, and improve data quality. For the multi-view gait images in the data, the original images are first normalized to ensure data consistency under different viewpoints and environmental conditions.

Gait analysis usually requires extracting human body contours from complex backgrounds to reduce the interference of irrelevant information on subsequent feature extraction [10, 11]. Assuming the input image is $\mathrm{I}(\mathrm{x},\mathrm{y})$ and the foreground mask after segmentation is $\mathrm{M}(\mathrm{x},\mathrm{y})$, the human body contour area $\mathrm{C}(\mathrm{x},\mathrm{y})$ can be expressed as:

$$\mathrm{C}(\mathrm{x},\mathrm{y}) = \mathrm{I}(\mathrm{x},\mathrm{y}) \cdot \mathrm{M}(\mathrm{x},\mathrm{y}) \tag{1}$$

$\mathrm{M}(\mathrm{x},\mathrm{y})$ is a binary mask used to mark the foreground area. This step can effectively remove background noise and retain the main contour information of the human body. In addition, in order to further improve the segmentation accuracy, morphological operations are introduced to optimize the segmentation results. The morphological opening operation of the segmentation mask is performed using the structural element S the formula is:

$$M_{opt}(x, y) = (M(x, y) \oplus S) \ominus S \tag{2}$$

$\oplus$ and $\ominus$ represent dilation and erosion operations, respectively. Through these operations, small holes and edge burrs generated during the segmentation process can be effectively eliminated, thereby improving the quality of the human body contour.

The data contains multi-view gait images. Due to the difference in perspective, the same action can show different spatial distributions under different perspectives, so it is necessary to align them spatially to unify the coordinate system [12, 13]. Assuming that the perspective transformation matrix is T_{ij}, the coordinates of point $P_i(x_i, y_i)$ in perspective i are converted to $P_j(x_j, y_j)$ in perspective j. The conversion formula is:

$$P_j = T_{ij} \cdot P_i \tag{3}$$

T_{ij} includes the rotation matrix R and the translation vector t, namely $T_{ij} = [R|t]$. In order to obtain the optimal T_{ij}, the method of minimizing the reprojection error is used for optimization. Specifically, by matching the key points in multiple perspectives, the objective function is constructed:

$$E(T_{ij}) = \sum_{k=1}^{K} \left\| P_j^k - T_{ij} \cdot P_i^k \right\|^2 \tag{4}$$

K is the number of matching points, P_i^k and P_j^k are the coordinates of key points in the perspectives i and j respectively. Through the iterative optimization algorithm, the optimal T_{ij} can be solved to ensure the spatial consistency of multi-view data [14,15].

In order to eliminate the scale inconsistency problem caused by individual height and weight differences, the gait image is scaled to a standard size using a normalization method. Assuming the height of the original image is H, the width is W, and the normalized size is H' and W', the normalized scale factor S is expressed as:

$$S = \min\left(\frac{H'}{H}, \frac{W'}{W}\right) \tag{5}$$

The normalized image pixel coordinate (x', y') calculation formula is:

$$x' = S \cdot x, y' = S \cdot y \tag{6}$$

Through the above steps, data preprocessing can improve the quality of the original image and lay the foundation for subsequent feature extraction. The processed image is shown in Figure 1.

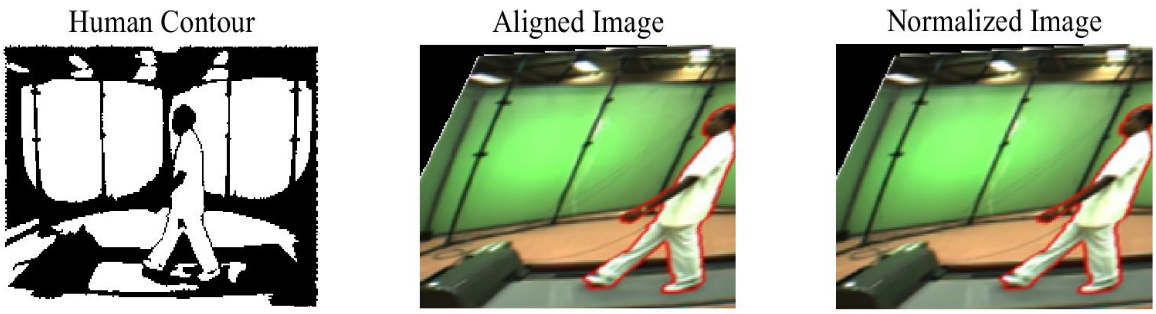

Fig. 1: Preprocessed image

Feature Extraction

Feature extraction aims to extract key gait features from human body contour images, such as joint positions, limb angles, and motion trajectories. This study uses CNN as the core tool and combines a multi-layer feature fusion strategy to achieve efficient and robust feature extraction [16, 17].

In CNN, the input image $I(x, y)$ is first normalized to a fixed size and converted to grayscale. Let the input image be $X \in R^{H \times W \times C}$, where H and W represent the height and width of the image, respectively, and C represents the number of channels.

Each layer of the convolution operation filters the input feature map through the convolution kernel W_l to extract local features. Assuming that the output feature map of the l layer is $F_l \in R^{H_l \times W_l \times C_l}$, the convolution formula is:

$$F_l = \sigma(W_l * F_{l-1} + b_l \quad (7)$$

W_l and b_l represent the convolution kernel weight and bias term, respectively, $*$ represents the convolution operation, and σ is the activation function used to introduce nonlinear characteristics. In each layer, the size, stride, and padding of the convolution kernel determine the spatial resolution of the feature map.

To reduce the dimension of the feature map and enhance the robustness to translation and scaling, a pooling layer is usually added after the convolution operation [18]. Assuming the pooling window size is $k \times k$, the pooling formula is:

$$F_{pool}(i,j) = \max_{m,n\in[0,k-1]} F_{comv}(i+m, j+n) \quad (8)$$

The pooling operation not only reduces the size of the feature map, but also retains important local features.

As the number of network layers increases, shallow features such as edges and textures are gradually transformed into deep features such as joint positions and limb shapes. These high-level features are crucial for gait analysis because they can capture the core patterns of human movement [19, 20].

Gait analysis needs to capture feature information of different scales, such as local joint motion and overall body posture, so multi-scale feature fusion is a key step. This paper uses Feature Pyramid Network (FPN) and Skip Connection to fuse scale features. FPN is used to extract multi-scale information from convolutional feature maps at different levels. Assuming that the feature map of the l layer is F_l and the upsampled feature map is U_l, the fusion formula is:

$$F_{fuse} = \sum_{l=1}^{L} \mathrm{Conv}(U_l) \quad (9)$$

Conv represents a 1x1 convolution operation, which is used to adjust the feature dimension. U_l is a low-resolution feature map obtained by upsampling the high-level feature map.

The jump connection combines the shallow features with the high-level features to retain more detailed information. Let the shallow feature map be F_{low} and the high-level feature map be F_{high}, then the jump connection formula is:

$$F_{combined} = F_{low} = U_p\mathrm{Sample}(F_{high}) \quad (10)$$

Multi-scale feature fusion can effectively cope with complex scenes in gait analysis. Shallow features can capture the movement details of local joints, while high-level features can reflect the posture changes of the whole body.

Key point localization is an important task in gait analysis. Its goal is to predict the key points of the human body, such as the position of the hip, knee and ankle joints, from the feature map.

The regression method is used to generate a heat map of key points. Each pixel value in the heat map represents the probability that the position belongs to the key point. Assuming the key point coordinates are x_k, y_k, the generation formula of the heat map H_k is:

$$H_k(x,y) = \exp\left(\frac{(x-x_k)^2+(y-y_k)^2}{2\sigma^2}\right) \quad (11)$$

σ controls the distribution range of the heat map. The peak position of the heat map is the predicted coordinate of the key point.

The formula for predicting the key point coordinates based on the maximum value position of the heat map is:

$$(x_k, y_k) = \arg\max_{(x,y)} H_k\,(x,y) \quad (12)$$

Since there is a fixed geometric relationship between human joints, the prediction results of key points can be further optimized through geometric constraints. Assuming the coordinates of the hip joint, knee joint, and ankle joint are (x_h, y_h), (x_k, y_k), and $(x_a,$ ya), respectively, the following constraints can be used for optimization:

$$x_k = \lambda x_h + (1-\lambda)x_a, y_k = \lambda y_h + (1-\lambda)y_a \quad (13)$$

λ is the proportional coefficient. Based on the key point coordinates, the angle of the limb can be further calculated. Assuming that the coordinates of the hip joint and knee joint are (x_1, y_1) and (x_2, y_2) respectively, the calculation formula for the limb angle θ is:

$$\theta = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \tag{14}$$

This angle information can reflect the dynamic changes in the gait cycle and provide rich feature support for subsequent modeling. The feature extraction and heat map are shown in Figure 2:

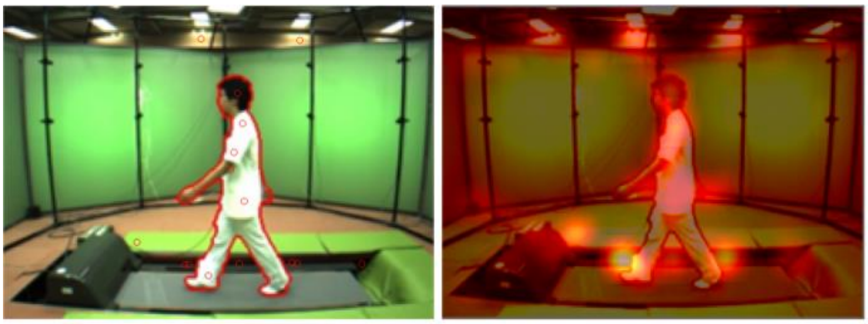

Fig. 2: Feature points and heatmaps

In the feature extraction process, convolutional feature extraction captures local and global features of the image through multi-layer convolution and pooling operations, and multi-scale feature fusion uses feature pyramid networks and jump connections to enhance the expressiveness of features. Key point positioning predicts the position of human joints through heatmap generation and geometric constraints, and calculates high-order features such as limb angles. These steps together constitute a complete feature extraction framework, providing high-quality input data for gait analysis.

Dynamic Modeling

The goal of dynamic modeling is to capture the temporal changes in the gait cycle. This study uses LSTM to model the gait feature sequence [21, 22].

LSTM can effectively capture long-term dependencies and complex temporal patterns through its unique gating mechanism [23]. The main concept is to use three gates (input, forget, and output gates) and cell state to address the issue of gradient vanishing or exploding in conventional recurrent neural networks in long sequence modeling.

The forget gate determines which information in the cell state needs to be discarded. Its calculation formula is:

$$f_t = \sigma(W_f \cdot [h_{t-1}, v_t] + b_f) \tag{15}$$

f_t is the output of the forget gate, W_f and b_f are the weight matrix and bias term of the forget gate respectively, σ is the Sigmoid activation function, which limits the output value to [0, 1], and $[h_{t-1}, v_t]$ represents the concatenation of the hidden state h_{t-1} at the previous moment and the current input v_t.

The output value f_t of the forget gate determines the degree of retention of each dimension of information in the cell state: close to 0 means that the information is completely forgotten, and close to 1 means that it is completely retained.

The input gate determines which new information can be written into the cell state and generates candidate cell states at the same time. Its calculation formula is:

$$i_t = \sigma(W_i \cdot [h_{t-1}, v_t] + b_i) \tag{16}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, v_t] + b_C) \tag{17}$$

Among Them

I_t is the output of the input gate, W_i and b_i are the weight matrix and bias term of the input gate, respectively, and $\tanh$ is the hyperbolic tangent activation function. The output value is limited to [-1, 1], $\tilde{C}_t$ is the candidate cell state, which represents the new information that may be added to the cell state.

DOI: 10.3844/ajbbsp.2026.22.02.024

The output value $\mathrm{i_t}$ of the input gate determines the degree of information written in the candidate cell state $\tilde{\mathrm{C}}_\mathrm{t}$.

The cell state $\mathrm{C_t}$ is the core component of LSTM, which is used to store long-term dependency information. Its update formula is:

$$\mathrm{C_t} = \mathrm{f_t} \odot \mathrm{C_{t-1}} + \mathrm{i_t} \odot \tilde{\mathrm{C}}_\mathrm{t} \tag{18}$$

Among Them

$\odot$ Represents element-by-element multiplication, $f_t \odot C_{t-1}$ represents the old cell state retained under the control of the forget gate, and $i_t \odot \tilde{C}_t$ represents the new information written under the control of the input gate. The update of the cell state ensures that LSTM can selectively retain or forget historical information while adding new relevant information.

The output gate determines the hidden state ht of the current time step, that is, the feature representation of the final output. Its calculation formula is:

$$\mathrm{o_t} = \sigma(\mathrm{W_o} \cdot [\mathrm{h_{t-1}}, \mathrm{v_t}] + \mathrm{b_o}) \tag{19}$$

$$\mathrm{h_t} = \mathrm{o_t} \odot \tanh(\mathrm{C_t}) \tag{20}$$

Among Them

$\mathrm{o_t}$ is the output of the output gate, $\mathrm{W_o}$ and $\mathrm{b_o}$ are the weight matrix and bias term of the output gate, respectively, $\tanh(\mathrm{C_t})$ represents the nonlinear transformation of the updated cell state, and $\mathrm{h_t}$ is the hidden state of the current time step.

The hidden state $\mathrm{h_t}$ is the output provided by LSTM to the outside, and also serves as one of the inputs of the next time step.

In gait analysis, LSTM is used to perform time series modeling on gait feature sequences to capture the dynamic changes in the gait cycle [24, 25]. Assuming that the gait cycle contains T time steps, the key point coordinates $(\mathrm{x_t}, \mathrm{y_t})$ and limb angles θ_t of each time step are concatenated into a feature vector $\mathrm{v_t}$, that is:

$$v_t = [x_t, y_t, \theta_t] \tag{21}$$

Assuming that the gait cycle contains T time steps, the gait feature sequence is represented as $V = [v_1, v_2, \dots v_T]$.

At each time step t, the input feature vector $\mathrm{v_t}$ and the hidden state h_{t-1} of the previous moment are input into the LSTM unit. Through the synergy of the forget gate, input gate and output gate, the cell state $\mathrm{C_t}$ and hidden state $\mathrm{h_t}$ are gradually updated. Finally, the hidden state h_T contains the timing information of the entire gait cycle [26] (Guan, 2025).

The hidden state h_T of the last time step is input into the fully connected layer, and the classification label y is generated through the Softmax function:

$$y = Softmax(W_y \cdot h_T + b_y) \tag{22}$$

W_y and b_y are the weight matrix and bias term of the fully connected layer, respectively.

Biomechanical Model Construction

Biomechanical model construction is a key step in associating gait characteristics with physical parameters. Based on the finite element analysis method, this paper combines imaging parameters and muscle dynamics indicators to construct a personalized joint dynamics model [27, 28].

Based on the imaging parameters, the finite element mesh of the human lower limbs is generated. Assume that the node displacement is u, the strain field is ϵ, the stress field is σ, the elastic matrix is D, and the finite element equilibrium equation is:

$$\mathrm{K} \cdot \mathrm{u} = \mathrm{F} \tag{23}$$

K is the stiffness matrix, and F is the external force vector. By solving this equation, the dynamic response of the joint is obtained.

Based on the principle of inverse dynamics, the joint torque τ is calculated. Assuming the ground reaction force is $\mathrm{F_{grf}}$ and the joint angle is γ, the torque formula is:

DOI: 10.3844/ajbbsp.2026.22.02.024

$$\tau = r \cdot F_{grf} \cdot \sin(\gamma) \tag{24}$$

r is the length of the lever arm. This formula is used to associate imaging parameters with muscle dynamics indices [29].

In view of individual anatomical differences, an iterative optimization algorithm is used to adjust the model parameters. Assuming the objective function is J, the optimization formula is:

$$\min_{p} J(p) = \sum_{i=1}^{N} \|y_i - \hat{y}_i(p)\|^2 \tag{25}$$

p is the model parameter, y_i and $\hat{y}_i$ are the true value and predicted value, respectively.

Parameter Optimization

Parameter optimization is an important link in the process of applying the gait analysis model. Its aim is to realize the minimum prediction error and enhance the generalization ability of the model by tuning the parameters of the model. In this paper, an iterative optimization algorithm is used to combine the loss function and regularization technology to associate the imaging parameter with muscle dynamic information so as to enhance the adaptability of the personalized biomechanical model.

Loss function is used to reflect the gap between the value that is estimated by the model and the true value and it is the core basis of the optimization algorithm. In this paper, an iterative optimization algorithm is used to combine loss function and regularization technology to associate the imaging parameter with muscle dynamic information so as to enhance the adaptability of personalized biomechanical model.

The loss function is used to measure the difference between the model's predicted value and the true value, and is the core basis of the optimization algorithm:

$$L = \frac{1}{N} \sum_{i=1}^{N} \|y_i - \hat{y}_i(p)\|^2 \tag{26}$$

Among them, N represents the number of samples, y_i and $\hat{y}_i$ represent the true value and predicted value of the i th sample respectively, and $\|y_i - \hat{y}_i(p)\|^2$ represents the square of the prediction error.

In addition, in order to comprehensively evaluate the model performance, the loss function can also be extended to the form of multi-task learning, including classification loss, regression loss, and regularization term:

$$L_{total} = L_{classification} + \lambda_1 L_{regression} + \lambda_2 R(p) \tag{27}$$

Among Them

$L_{classification}$ represents the cross-entropy loss of gait classification, $L_{regression}$ represents the mean square error of continuous variables such as joint angle and plantar pressure, $R(p)$ represents the regularization term, and λ_1 and λ_2 are hyperparameters used to balance various losses.

In order to prevent the model from overfitting, regularization technology is introduced to limit the complexity of model parameters. The definition of the regularization term $R(p)$ is:

$$R(p) = \frac{1}{2} \|p\|^2 \tag{28}$$

$\|p\|^2$ represents the sum of squares of model parameters.

After the total loss function includes the regularization term, the parameter update formula becomes:

$$p_{t+1} = p_t - \mu \cdot (\nabla_p L + \lambda \cdot p_t) \tag{29}$$

λ is the regularization coefficient, which controls the regularization strength.

Stochastic Gradient Descent (SGD) is used to optimize the model parameters, and the parameters are updated by randomly sampling small batches of data. The update formula is:

$$p_{t+1} = p_t - \mu \cdot \nabla_p L_{batch} \tag{30}$$

Lbatch represents the loss function of small batch data.

Definition of parameter optimization in gait analysis: In gait analysis, connect parameters from the imaging process with muscle dynamic indicators and therefore create an individual biomechanical model [30, 31]. Random initialization of a pretrained model parameters can be used as initial values for a fast convergence of the optimization process. Joint optimization objectives can be used for the different tasks: gait classification, joint angle prediction, and plantar pressure distribution. The joint optimization formula is:

$$\mathrm{L_{joint}} = \alpha_1 \mathrm{L_{classification}} + \alpha_2 \mathrm{L_{angle}} + \alpha_3 \mathrm{L_{pressure}} \tag{31}$$

Among them, $\alpha_1, \alpha_2, \alpha_3$ are the task weights. In gait analysis, parameter optimization solves the problem that traditional methods are difficult to adapt to individual anatomical differences, and provides reliable technical support for personalized biomechanical modeling.

Experimental Design

This paper is to thoroughly validate the performance of the method on the task of gait classification and its application to biomechanical parameter analysis and personalized modeling. The experimental data are extracted from the public OU-ISIR gait database [32], which is composed of gait images from multiple views, namely OU-MVLP and Clothing subsets. These data have been collected by the Institute of Scientific and Industrial Research (ISIR) since 2007. Different from other public gait data sets, the data are high-resolution videos (640x480 pixels) recorded at 60 fps. Moreover, it records gait information of different persons on the treadmill. Therefore, the dataset also provides rich covariate conditions, such as clothes change, perspective difference, etc., which offer the possibility to verify the robustness and generalization performance of the model:

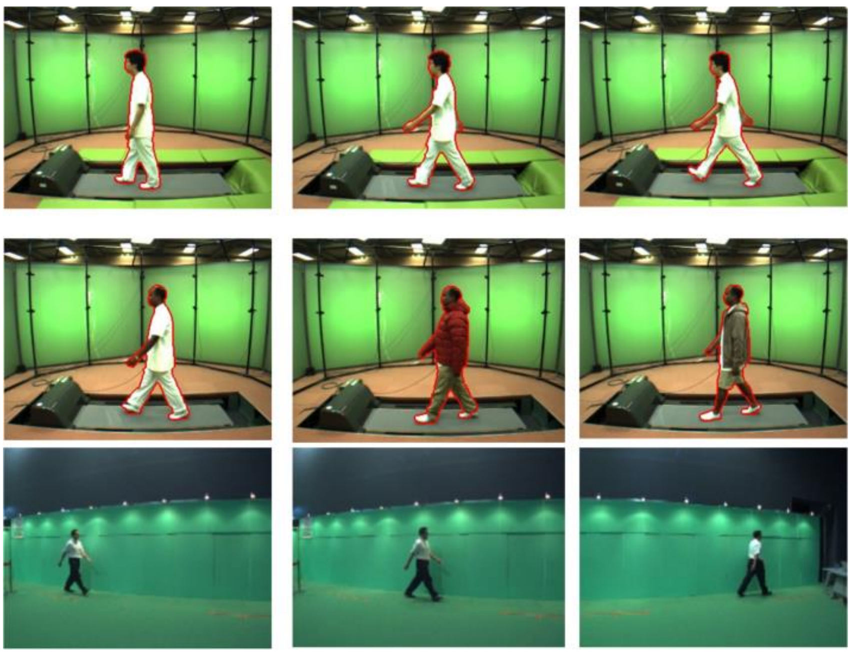

Fig. 3: Some dataset images

In order to validate the effectiveness of the method, the paper selects a number of representative algorithms for a comprehensive assessment of the proposed method. Firstly, it selects the traditional gait analysis method based on markers. Compared with other devices, the Xsens inertial sensor system has higher accuracy and reliability in a laboratory environment, and it is used as the benchmark of basic accuracy of the method. Secondly, the paper introduces markerless OpenPose and Theia3D method to validate the adaptability of the equipment with low cost and multi-view modeling ability to the complex scene. In addition, the paper compares two advanced models in deep learning field, CNN-LSTM hybrid model and CNN-CapsNet capsule network model. The former is good at capturing the spatiotemporal change of gait features, and the latter solves the covariate problem of clothing change and perspective difference, which can play a role in highlighting the advantages of the method in dynamic modeling and personalized adaptation. Through the comparison with above specific algorithms, it is shown that the accuracy, generalization ability and practical application value of proposed method are demonstrated in multiple aspects.

DOI: 10.3844/ajbbsp.2026.22.02.024

In terms of comparison indicators, multiple quantitative indicators are selected to measure model performance, including gait classification accuracy, joint angle error, plantar pressure distribution deviation, ground reaction force error, and model adaptation time. These indicators were selected based on their importance and representativeness in gait analysis. Gait classification accuracy can reflect the robustness of the model under different covariate conditions. Joint angle error and plantar pressure distribution deviation are directly related to the analytical accuracy of biomechanical parameters. Ground reaction force error is used to evaluate the model's ability to predict the dynamic gait cycle. Model adaptation time reflects the practical application value of the method, especially in scenarios with high real-time requirements.

Results Analysis

Classification Performance Analysis

In order to evaluate the classification performance of different methods under covariate conditions, clothing changes and perspective differences, the classification accuracy of six gait analysis methods was compared. Table 1 shows the classification accuracy of each method under clothing changes and perspective differences.

Table 1: Classification accuracy under clothing changes and perspective differences

Method	Costume change accuracy (%)	Viewing angle difference accuracy (%)
Xsens inertial sensor	95.3	87.6
OpenPose	92.8	80.4
Theia3D	94.1	83.7
CNN-LSTM	96.2	89.1
CNN-CapsNet	95.8	88.5
Method in this paper	97.5	91.2

As can be seen from Table 1, under the condition of clothing changes, the classification accuracy of this method reaches 97.5%, which is significantly higher than other methods. The Xsens inertial sensor system and CNN-LSTM model also show high classification accuracy, 95.3% and 96.2%, respectively, but there is still a gap compared with this method. Open Pose and Theia3D have relatively low accuracies of 92.8% and 94.1%, respectively, indicating their limited adaptability to clothing changes. Under the condition of view difference, the paper's method also performs best with a classification accuracy of 91.2%, while the accuracies of Open Pose and Theia3D drop to 80.4% and 83.7%, respectively. This result shows that the proposed method can effectively reduce the classification error caused by clothing changes or perspective differences by integrating multimodal image analysis and sports biomechanics modeling, showing stronger robustness. In addition to classification accuracy, classification error in multi-view data is also a key indicator for evaluating model performance. Table 2 further refines the average classification error, maximum error perspective, and minimum error perspective of different methods in multi-view data to fully reflect the performance of the model under different view conditions.

Table 2: Classification error in multi-view data

Method	Average classification error (%)	Maximum error viewing angle (°)	Minimum error viewing angle (°)
Xsens inertial sensor	8.5	45	0
OpenPose	11.2	60	15
Theia3D	9.7	55	10
CNN-LSTM	7.8	50	5
CNN-CapsNet	8.2	52	8
Methods in this paper	6.3	40	2

From the data in Table 2, the average classification error of the paper's method in multi-view data is the lowest, only 6.3%, which is significantly better than other methods. The average error of the Xsens inertial sensor system is 8.5%, while the errors of OpenPose and Theia3D are 11.2% and 9.7%, respectively. In addition, the maximum error viewing angle of the paper's method is 40°, which is much lower than 60° of Open Pose and 55° of Theia3D, indicating that it is more stable at extreme viewing angles. In terms of the minimum error viewing angle, the paper's method achieves the best performance at 2°, while the best viewing angle range of other methods is usually above 5°. These results further verify the superiority of the paper's method in multi-view gait classification, especially its adaptability under complex viewing conditions. Overall, the paper's method performs well in average error and can maintain high classification accuracy under extreme viewing angles, which fully demonstrates its potential in practical applications.

There are significant differences in the classification performance of different methods under covariate conditions. As can be seen from Table 1, the classification accuracy of the proposed method is the highest under the conditions of clothing changes and perspective differences, which are 97.5% and 91.2% respectively. This is due to its ability to integrate multimodal image analysis with motion biomechanical modeling, which effectively reduces the errors caused by clothing and perspective changes. Table 2 further shows that the proposed method has the lowest average classification error in multi-view data, and the maximum error and minimum error view ranges are smaller, reflecting its robustness and stability under complex view angles. These excellent performances are mainly attributed to the ability of CNN to extract key gait features and LSTM to model the timing of dynamic gait cycles, which enables the model to better adapt to changing environments. However, other methods, such as Open Pose and Theia3D, rely on visual input or skeleton reconstruction. The error is large at extreme viewing angles, and although the Xsens inertial sensor has the smallest error at 0° viewing angle, its adaptability to viewing angle changes is limited.

Evaluation of Biomechanical Parameter Accuracy

Joint angle error is an important indicator for evaluating the accuracy of biomechanical parameter analysis of gait analysis methods, and directly reflects the accuracy of the model in capturing human motion characteristics. In order to compare the performance of different methods in joint angle prediction, the error distribution of each method on the hip joint, knee joint, and ankle joint is shown in Figure 4:

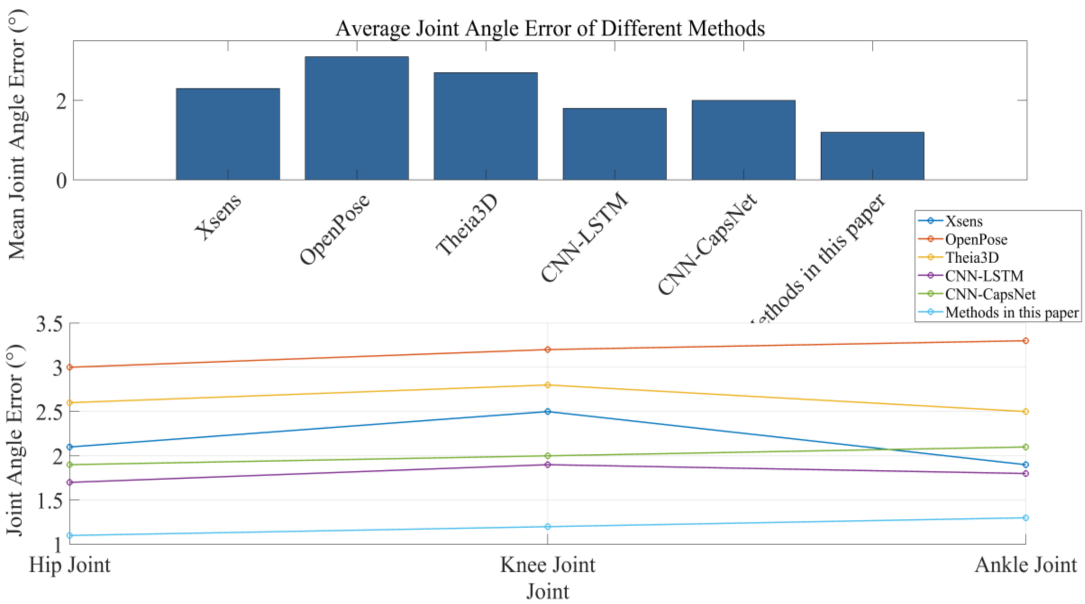

Fig. 4: Distribution of joint angle errors of different methods

As can be seen from Figure 4, the angle errors of the proposed method at the hip, knee, and ankle joints are significantly lower than those of other methods. The error at the hip joint is only 1.1°, while the errors of the Xsens inertial sensor and Open Pose are 2.1° and 3.0°, respectively. This shows that the proposed method can effectively reduce the errors caused by clothing changes or perspective differences by extracting key gait features through CNN and combining LSTM to model the dynamic gait cycle. In addition, although the errors of CNN-CapsNet and CNN-LSTM in the knee and ankle joints are small, their overall performance is still inferior to that of the proposed method, and the gap in the error of the hip joint is obvious. This result verifies the high accuracy and robustness of the proposed method in joint angle prediction, especially its adaptability under complex covariate conditions.

The deviation of plantar pressure distribution is a key indicator to measure the accuracy of ground reaction force prediction by gait analysis methods, which directly affects the analytical quality of dynamic load distribution in the gait cycle. In order to comprehensively evaluate the performance of different methods in predicting plantar pressure distribution, the error distribution of each method in the forefoot, arch, and heel area was statistically analyzed and compared with the overall root mean square error.

The results are shown in Figure 5:

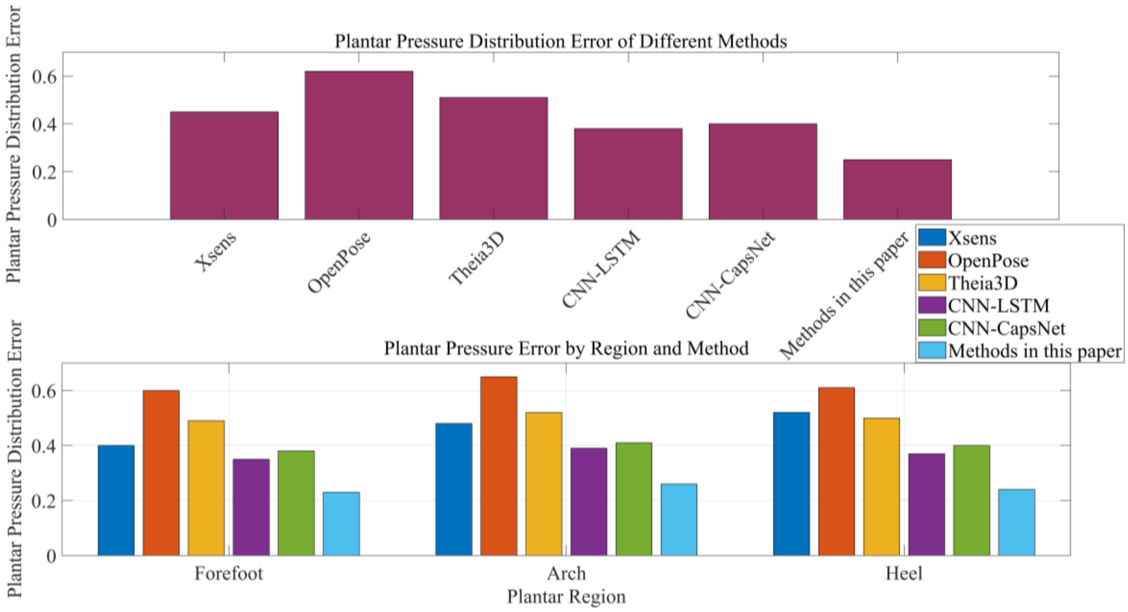

Fig. 5: Comparison of plantar pressure distribution deviation of different methods

As can be seen from Figure 5, the proposed method shows significant advantages in plantar pressure distribution deviation. In the forefoot area, the RMSE of the proposed method is only 0.23, which is much lower than 0.60 of Open Pose and 0.50 of Theia3D. This is mainly due to the fact that the proposed method can more accurately capture the dynamic changes of plantar pressure by integrating multimodal images and sports biomechanics modeling. In addition, CNN-Caps Net and CNN-LSTM perform relatively well in the arch and heel areas, but the deviation is still higher than that of the proposed method. Especially in the heel area, the error of the proposed method is only 0.24. This result further verifies the high accuracy and stability of the proposed method in predicting plantar pressure distribution and provides reliable technical support for gait analysis.

Model Adaptability Analysis

Ground Reaction Force (GRF) is a core indicator in gait analysis, which can reflect the changes in dynamic load distribution during the gait cycle. In order to evaluate the performance of different methods in adapting to various gait abilities, Table 3 shows the GRF error distribution of each method under different gait abilities, slow walking, normal walking, fast walking, and running.

Table 3: Ground reaction force under different gaits

Method	Slow walking (%)	Normal walking (%)	Fast walking (%)	Running (%)	Average error (%)
Xsens inertial sensor	12.5	10.8	14.2	16.7	13.55
OpenPose	15.4	13.7	17.6	20.1	16.7
Theia3D	14.1	12.9	15.8	18.3	15.28
CNN-LSTM	9.8	8.5	11.4	13.2	10.73
CNN-CapsNet	10.2	9.1	12.3	14.5	11.53
Method in this paper	8.7	7.9	10.1	11.8	9.63

DOI: 10.3844/ajbbsp.2026.22.02.024

As can be seen from Table 3, under slow walking conditions, the error of this method is only 8.7%, while the errors of the Xsens inertial sensor and Open Pose are 12.5% and 15.4%, respectively. This shows that this method can effectively capture the dynamic load changes in low-speed gait and reduce the error caused by gait speed changes. Under high-speed gait conditions such as running, the error of this method is 11.8%, which is much lower than 20.1% of Open Pose and 18.3% of Theia3D, indicating that this method has higher prediction accuracy in complex dynamic gaits.

Furthermore, the average GRF error of this method is 9.63%, and it is obviously better than other methods' error. The main reason for this phenomenon is that the paper's method can fully adapt to individual anatomical differences and dynamic load distribution under different gait abilities based on motion biomechanical finite element analysis and multimodal image feature extraction when the gait ability varies greatly. While, because of the lack of personalized model ability, the error of Xsens inertial sensors and Open Pose can increase a lot when the gait ability varies greatly.

This paper believed that the capability of adapting to individual anatomical variation is an important metric for evaluating the performance of personalized modeling of gait analysis models, and the model adaptability to various populations depends on the performance. Therefore, the adaptability scores of all methods on the above anatomical parameters, including height, weight, leg length, and stride, are shown in Figure 6.

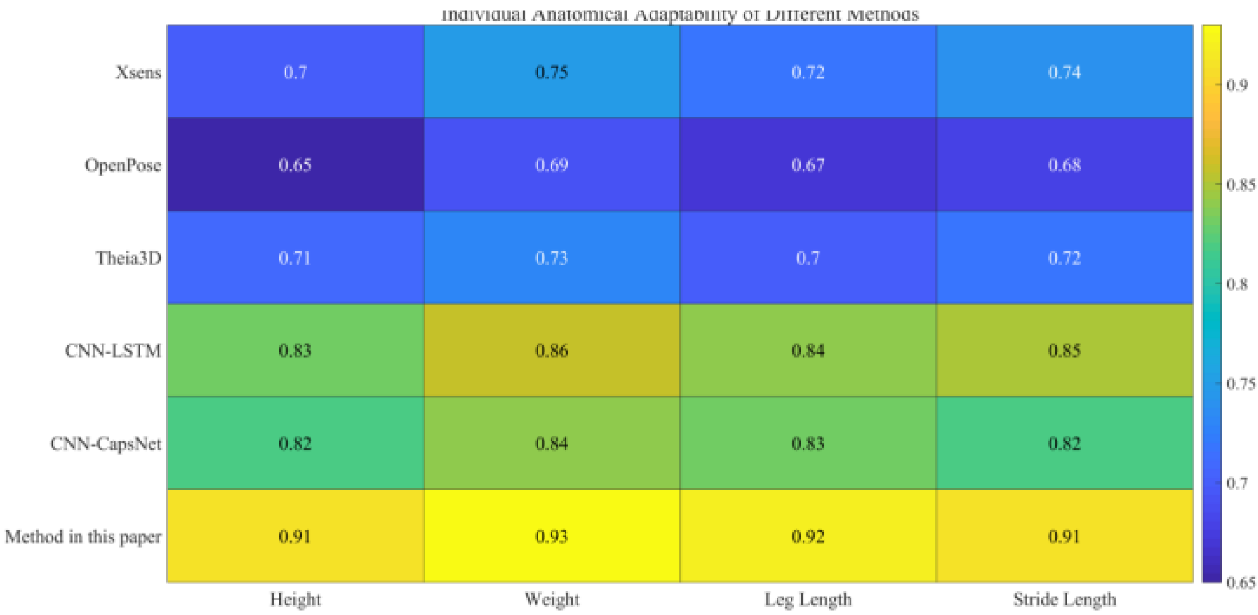

Fig. 6: Adaptability scores on key parameters

As can be seen from Figure 6, the proposed method performs best in adapting to individual anatomical differences, with an overall score of about 0.92, significantly higher than other methods. In terms of height parameters, the proposed method scores 0.91, while the Xsens inertial sensor system and Open Pose scores are 0.70 and 0.65, respectively, indicating that the proposed method can better adapt to individual differences of different heights. Similarly, in terms of weight and leg length parameters, the scores of the paper's method are much higher than those of Theia3D and CNN-LSTM, reaching 0.93 and 0.92 respectively. This advantage is mainly attributed to the fact that the paper's method associates imaging parameters with muscle dynamics indicators through an iterative optimization algorithm, thereby achieving accurate adaptation to individual anatomical differences. In addition, in terms of stride parameters, the score of this method is 0.91, which is significantly higher than other algorithms, further verifying its superiority in personalized modeling. Overall, this method has stronger adaptability and higher prediction accuracy when dealing with individual anatomical differences.

Motion-Image Synchronization Error and Long-Term Stability Test

Motion-image synchronization error is an important indicator to measure the accuracy of multimodal data fusion. Especially in gait analysis, it is crucial to ensure the temporal consistency of motion capture data and image data. This study compares the motion sensor time and image frame timestamp, calculates the deviation distribution of the two, and introduces multiple indicators to comprehensively evaluate the synchronization performance. The results are shown in Table 4.

As can be seen from Table 4, the proposed method performs best in all indicators of motion-image synchronization error. The average synchronization error is only 5.3 ms, which is significantly lower than other methods. At the same time, the standard deviation of the time deviation is 2.4 ms, indicating that its synchronization error distribution is more concentrated and stable. The frame alignment success rate reaches 98.6%, which is much higher than other methods, further verifying the high-precision time synchronization capability. In addition, the event delay fluctuation range is the smallest, only [1.5, 3.2] ms, which reflects the robustness and adaptability of the system in dynamic environments. In contrast, although CNN-Caps Net performs better, its average synchronization error (22.4 ms) and frame alignment success rate (91.8%) are still significantly inferior to those of the proposed method, indicating that the proposed method has significant advantages in temporal consistency in multimodal data fusion.

The long-term stability test is designed to evaluate the performance changes of the system under continuous operation conditions. Especially in clinical monitoring or smart device application scenarios, the stability of the system directly affects the reliability of user experience and data analysis. After 72 hours of continuous operation, this paper compares the changes in classification accuracy, joint angle error, ground reaction force error, and hardware resource usage of each method. The results are shown in Table 5.

Table 4: Motion-image synchronization error of different methods

Method	Average synchronization error (ms)	Maximum synchronization error (ms)	Minimum synchronization error (ms)	Time deviation standard deviation (ms)	Frame alignment success rate (%)	Event delay fluctuation range (ms)
Xsens inertial sensor	25.6	48.2	8.5	7.3	90.2	[6.2, 10.8]
OpenPose	32.8	56.7	10.4	9.1	86.5	[8.7, 15.4]
Theia3D	28.3	50.9	9.8	8.6	88.7	[7.5, 12.3]
CNN-LSTM	18.5	35.4	6.2	5.8	93.1	[4.8, 8.2]
CNN-CapsNet	22.4	45.6	7.8	6.5	91.8	[5.2, 10.1]
Method in this paper	5.3	12.1	2.8	2.4	98.6	[1.5, 3.2]

Table 5: Long-term stability test results

Method	Classification accuracy decreased (%)	Joint angle error increase (°)	GRF error increase (%)	CPU usage change (%)	Memory usage change (MB)
Xsens inertial sensor	5.2	0.8	8.4	12.3	185
OpenPose	7.6	1.2	10.3	15.6	210
Theia3D	6.5	1	9.1	14.2	200
CNN-LSTM	3.8	0.5	6.7	8.5	150
CNN-CapsNet	4.2	0.7	7.3	9.8	165
Method in this paper	1.5	0.2	3.1	3.2	90

As can be seen from Table 5, the performance of this method in the long-term stability test is significantly better than that of other methods. The classification accuracy drops by only 1.5%, which is much lower than that of Xsens inertial sensors and OpenPose, indicating that its performance degradation is minimal under continuous operation conditions. At the same time, the joint angle error increased by only 0.2°, and the GRF error increased by 3.1%, which were significantly lower than other methods, reflecting its stability and reliability in predicting dynamic mechanical indicators. In addition, the CPU usage of the proposed method changed by +3.2%, and the memory usage changed by +90 MB, and the increase in resource consumption was also the smallest, further verifying its advantages in hardware compatibility and resource utilization. In contrast, although CNN-LSTM outperforms other methods, it is still inferior to the proposed method in terms of decreased classification accuracy, joint angle error, and resource usage. Overall, the proposed method shows higher stability and efficiency under long-term operation conditions, which provides strong support for its continued reliability in practical applications.

DOI: 10.3844/ajbbsp.2026.22.02.024

Time Efficiency Optimization

Model adaptation time is an important indicator to measure the practical application value of gait analysis methods. Especially in personalized modeling, building an efficient and accurate biomechanical model is crucial to improving user experience. In order to compare the model adaptation time performance of different methods when the number of samples increases, Figure 7 shows the distribution of adaptation time of each method under different sample numbers.

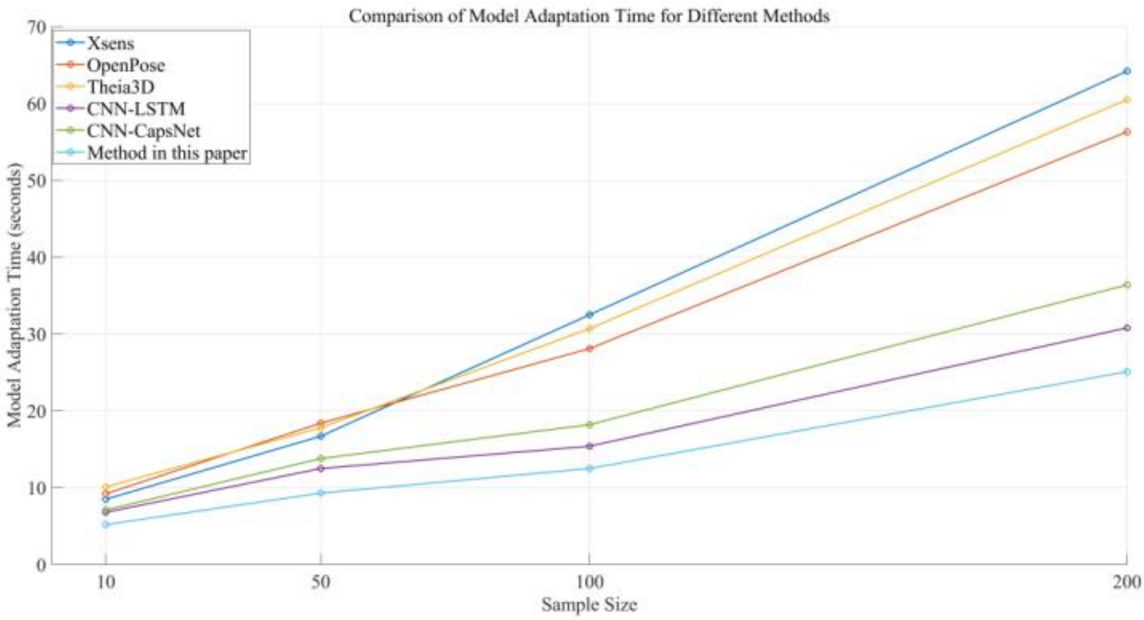

Fig. 7: Fitting time under different sample numbers

As can be seen from Figure 7, the model fitting time of the paper's method under different sample numbers is significantly lower than that of other methods. For example, with 10 samples, the fitting time of the paper's method is 5.2 seconds, while the fitting time of the Xsens inertial sensor system and Open Pose is 8.5 seconds and 9.2 seconds, respectively. This shows that the proposed method can quickly complete personalized modeling by optimizing the algorithm flow and reducing redundant calculations. In addition, with 200 samples, the adaptation time of the proposed method is only 25.1 seconds, which is much lower than the 60.5 seconds of Theia3D and the 30.8 seconds of CNN-LSTM, indicating that it can still maintain high efficiency when processing large-scale data.

Further observation shows that as the number of samples increases, the adaptation time of all methods shows a linear or nearly linear growth trend. However, the growth slope of the proposed method is the lowest, indicating that it has stronger scalability and adaptability in large-scale data scenarios. In contrast, traditional methods such as Xsens inertial sensors and OpenPose have a significant increase in adaptation time when the number of samples increases due to the lack of efficient optimization strategies.

Real-time performance is a key indicator for evaluating the applicability of gait analysis methods in dynamic environments, and directly affects the model's deployment capabilities in clinical monitoring and smart devices. In order to comprehensively evaluate the real-time performance of different methods, the real-time scores of each method in the key steps of data preprocessing, feature extraction, dynamic modeling, and model optimization were analyzed, and the results are shown in Figure 8.

As shown in Figure 8, the real-time scores of this method in the processing steps are better than those of other methods. When it comes to the data preprocessing stage, the score of this method is 0.91. Compared with 0.6 of the Xsens inertial sensor system and 0.68 of Open Pose, the advantages of the image segmentation and normalization algorithm are very obvious. Similarly, when it comes to the dynamic modeling stage, the score of the method in the paper is 0.92, which is much higher than 0.68 for Theia3D and 0.84 for CNN-LSTM. It shows that the computational efficiency in reflecting the dynamic changes of the gait cycle is higher. At last, the overall real-time score of this method is about 0.92, which verifies that the

DOI: 10.3844/ajbbsp.2026.22.02.024

method in the paper is applicable and efficient in a dynamic environment. From the above results, it can be seen that the proposed method can complete the model adjustment quickly and has good performance in a dynamic environment with high requirements for real-time. The technical support provides for the practical application of gait analysis.

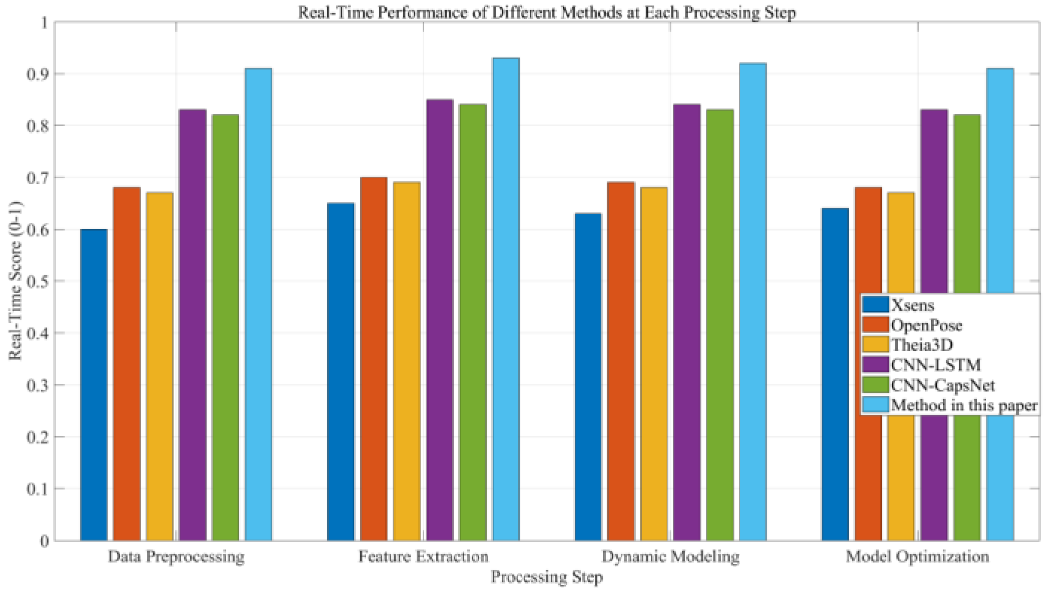

Fig. 8: Real-time comparison of different methods

Ablation Experiment

In order to further verify the effectiveness of each module in this method, an ablation experiment was designed to gradually remove key modules and compare their performance in indicators such as classification performance, biomechanical parameter analysis accuracy, model adaptability, and time efficiency. Through this method, the contribution of each module to the overall performance is quantitatively evaluated. The ablation experiment results are shown in Table 6.

Table 6: Ablation experiment results

Method	Classification accuracy (%)	Joint angle error (°)	Plantar pressure distribution deviation (RMSE)	Ground reaction force error (%)	Individual anatomical differences in fitness score	Model fitting time (seconds/sample)
Full Model	97.5	1.1	0.23	9.63	0.92	5.2
CNN only	92.5	2.8	0.55	15.8	0.75	15.8
No LSTM	94.8	2.3	0.48	12.3	0.82	12.5
No Personalized Modeling	96	1.8	0.35	10.5	0.88	10.8
No Parameter Optimization	95.5	2	0.38	11.2	0.85	9.5

From the ablation experiment results in Table 6, it can be seen that each module in this method has a significant contribution to the overall performance. The complete model performs best in key indicators such as classification accuracy, joint angle error, plantar pressure distribution deviation, ground reaction force error, and model adaptation time, indicating that the synergy between modules can significantly improve the accuracy and efficiency of gait analysis. After removing dynamic modeling or personalized modeling, the classification accuracy dropped to 94.8% and 96.0%, respectively, and the joint angle error and plantar pressure distribution deviation also increased significantly. This shows the importance of LSTM dynamic modeling and personalized biomechanical models in capturing gait timing changes and adapting to individual anatomical differences.

In addition, the lack of parameter optimization module leads to a further decline in various performance indicators, with the classification accuracy dropping to 95.5% and the ground reaction force error rising to 11.2%, which shows that the iterative optimization algorithm plays a key role in improving the prediction accuracy and generalization ability of the model. Overall, the method in this paper performs well under complex covariate conditions through the organic combination of multiple modules, and can also maintain high efficiency in large-scale data scenarios, which fully verifies its design rationality and practical application value.

Discussion

This study proposes a gait assessment method that integrates sports biomechanics and multimodal image analysis, significantly outperforming existing mainstream technologies across multiple dimensions. Compared to traditional inertial sensor systems relying on marker points or label-free methods based on single-view vision, our method effectively mitigates covariate interference caused by clothing variations and viewpoint differences through a CNN-LSTM joint architecture, achieving classification accuracies of 97.5% and 91.2%, respectively. This improvement stems not only from deep learning's robust extraction of contour features but also from dynamic temporal modeling's capture of the inherent regularity of gait cycles, thus exhibiting stronger generalization performance in complex real-world scenarios.

Regarding biomechanical parameter analysis, this paper introduces finite element analysis and personalized joint dynamics modeling, significantly improving the accuracy of joint angle prediction and the quality of plantar pressure distribution reconstruction. This demonstrates that coupling anatomical information from imaging with musculoskeletal dynamics indices through iterative optimization can effectively compensate for the impact of soft tissue artifacts and individual morphological differences on motion estimation. In contrast, while purely data-driven models can fit apparent movement patterns, they struggle to guarantee biomechanical validity, and traditional inverse dynamics methods are limited by general model assumptions and cannot adapt to individual specificities. Our proposed method strikes a balance between accuracy and physiological interpretability.

Notably, ablation experiments further validated the necessity of each module: removing LSTM led to a decrease in temporal modeling ability and a 2.7 percentage point reduction in classification accuracy; canceling personalized modeling caused the anatomical fit score to drop from 0.92 to 0.88, particularly showing degradation in key parameters such as leg length and stride length. Furthermore, the parameter optimization strategy not only improved prediction consistency but also reduced the single-sample model fit time to 5.2 seconds, significantly outperforming the comparative methods. This provides feasible support for rapid clinical deployment and real-time feedback, particularly suitable for applications with high timeliness requirements such as rehabilitation assessment and intelligent prosthetic adjustment.

Although this study achieved good results on the publicly available OU-ISIR dataset, several limitations remain. The current model relies on high-quality contour image input, which may degrade performance under extreme occlusion or low-light conditions; additionally, the construction of the finite element model still requires certain prior anatomical knowledge and has not yet achieved full end-to-end automation. Future work will explore end-to-end differentiable biomechanical networks that combine 3D reconstruction with physical constraints, and extend them to pathological gait populations to verify their universal value in clinical diagnosis.

Conclusion

This study successfully achieved its research goal: proposing a gait assessment method integrating sports biomechanics and multimodal imaging analysis to overcome marker positioning error, soft tissue interference, and insufficient personalized modeling in traditional gait analysis. The CNN-LSTM framework extracts robust gait features and models dynamic gait cycles, reducing interference from clothing changes and perspective differences. The personalized joint dynamics model built via finite element analysis, combined with iterative optimization, effectively links imaging parameters with muscle dynamics indicators and improves adaptability to individual anatomical differences. Tests on the OU-ISIR gait database show that the method achieves 97.5% classification accuracy under clothing changes and 91.2% under perspective differences. The joint angle prediction errors of hip, knee, and ankle are 1.1, 1.2, and 1.3, respectively, outperforming existing methods. It also performs better in plantar pressure estimation, ground reaction force prediction, individual adaptability, synchronization precision, and long-term stability. Ablation experiments verify the effectiveness of each module.

Acknowledgment

The author acknowledges the support and contributions that facilitated the completion of this work.

Funding Information

This research received no external funding.

Ethics

This study does not involve ethical review or approval.

References

1. Feng, Y. (2024). A study on the test-retest reliability of the Xsens inertial sensor system in gait analysis. Leather Science and Engineering, 34(5), 81–85.
2. Sethi, D., Bharti, S., & Prakash, C. (2022). A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work. Artificial Intelligence in Medicine, 129, 102314. https://doi.org/10.1016/j.artmed.2022.102314
3. Stenum, J., Hsu, M. M., Pantelyat, A. Y., & Roemmich, R. T. (2024). Clinical gait analysis using video-based pose estimation: Multiple perspectives, clinical populations, and measuring change. PLOS Digital Health, 3(3), e0000467. https://doi.org/10.1371/journal.pdig.0000467
4. Wren, T. A. L., Isakov, P., & Rethlefsen, S. A. (2023). Comparison of kinematics between Theia markerless and conventional marker-based gait analysis in clinical patients. Gait & Posture, 104, 9–14. https://doi.org/10.1016/j.gaitpost.2023.05.029
5. Semwal, V. B., Kim, Y., Bijalwan, V., Verma, A., Singh, G., Gaud, N., Baek, H., & Khan, A. M. (2023a). Development of the LSTM Model and Universal Polynomial Equation for All the Sub-Phases of Human Gait. IEEE Sensors Journal, 23(14), 15892–15900. https://doi.org/10.1109/jsen.2023.3281401
6. Vidya, B., & Perumal, S. (2022). Parkinson's disease diagnosis and stage prediction based on gait signal analysis using EMD and CNN–LSTM network. Engineering Applications of Artificial Intelligence, 114, 105099. https://doi.org/10.1016/j.engappai.2022.105099
7. Berke Erdaş, Ç., Sümer, E., & Kibaroğlu, S. (2022). CNN-based severity prediction of neurodegenerative diseases using gait data. Digital Health, 8, 205520762210751. https://doi.org/10.1177/20552076221075147
8. Merlin Linda, G., Sree Rathna Lakshmi, N. V. S., Murugan, N. S., Mahapatra, R. P., Muthukumaran, V., & Sivaram, M. (2022). Intelligent recognition system for viewpoint variations on gait and speech using CNN-CapsNet. International Journal of Intelligent Computing and Cybernetics, 15(3), 363–382. https://doi.org/10.1108/ijicc-08-2021-0178
9. Li, J., Liang, W., Yin, X., Li, J., & Guan, W. (2023). Multimodal Gait Abnormality Recognition Using a Convolutional Neural Network–Bidirectional Long Short-Term Memory (CNN-BiLSTM) Network Based on Multi-Sensor Data Fusion. Sensors, 23(22), 9101. https://doi.org/10.3390/s23229101
10. Bonanno, M., De Nunzio, A. M., Quartarone, A., Militi, A., Petralito, F., & Calabrò, R. S. (2023). Gait Analysis in Neurorehabilitation: From Research to Clinical Practice. Bioengineering, 10(7), 785. https://doi.org/10.3390/bioengineering10070785
11. Moro, M., Marchesi, G., Hesse, F., Odone, F., & Casadio, M. (2022). Markerless vs. Marker-Based Gait Analysis: A Proof of Concept Study. Sensors, 22(5), 2011. https://doi.org/10.3390/s22052011
12. Sarkar, M. (2025). Integrating Machine Learning and Deep Learning Techniques for Advanced Alzheimer's Disease Detection through Gait Analysis. Journal of Business and Management Studies, 7(1), 140–147. https://doi.org/10.32996/jbms.2025.7.1.8
13. Trabassi, D., Serrao, M., Varrecchia, T., Ranavolo, A., Coppola, G., De Icco, R., Tassorelli, C., & Castiglia, S. F. (2022). Machine Learning Approach to Support the Detection of Parkinson's Disease in IMU-Based Gait Analysis. Sensors, 22(10), 3700. https://doi.org/10.3390/s22103700
14. Li, W., Zou, K., Guo, J., Zhang, C., Feng, J., You, J., Cheng, G., Zhou, Q., Kong, M., Li, G., Guo, C. F., & Yang, J. (2024). Integrated Fibrous Iontronic Pressure Sensors with High Sensitivity and Reliability for Human Plantar Pressure and Gait Analysis. ACS Nano, 18(22), 14672–14684. https://doi.org/10.1021/acsnano.4c02919
15. Yamamoto, M., Shimatani, K., Ishige, Y., & Takemura, H. (2022). Verification of gait analysis method fusing camera-based pose estimation and an IMU sensor in various gait conditions. Scientific Reports, 12(1), 17719. https://doi.org/10.1038/s41598-022-22246-5
16. Kumar, D., & Verma, R. (2024). Human Gait Analysis with Hybrid Deep Learning Network CNN-SVM. 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), 482–486. https://doi.org/10.1109/icssas64001.2024.10760563
17. Erdaş, Ç. B., & Sümer, E. (2024). CNN-Based Neurodegenerative Disease Classification Using QR-Represented Gait Data. Brain and Behavior, 14(10), e70100. https://doi.org/10.1002/brb3.70100
18. Arshad, M. Z., Jamsrandorj, A., Kim, J., & Mun, K.-R. (2022). Gait Events Prediction Using Hybrid CNN-RNN-Based Deep Learning Models through a Single Waist-Worn Wearable Sensor. Sensors, 22(21), 8226. https://doi.org/10.3390/s22218226
19. Mogan, J. N., Lee, C. P., Lim, K. M., Ali, M., & Alqahtani, A. (2023). Gait-CNN-ViT: Multi-Model Gait Recognition with Convolutional Neural Networks and Vision Transformer. Sensors, 23(8), 3809. https://doi.org/10.3390/s23083809

20. Mehmood, A., Khan, M. A., Sharif, M., Khan, S. A., Shaheen, M., Saba, T., ... & Ashraf, I. (2024). Prosperous human gait recognition: An end-to-end system based on pre-trained CNN features selection. Multimedia Tools and Applications, 83(5), 14979-14999. https://doi.org/10.1007/s11042-020-08928-0
21. Khokhlova, M., Migniot, C., Morozov, A., Sushkova, O., & Dipanda, A. (2019). Normal and pathological gait classification LSTM model. Artificial Intelligence in Medicine, 94, 54–66. https://doi.org/10.1016/j.artmed.2018.12.007
22. Sarshar, M., Polturi, S., & Schega, L. (2021). Gait Phase Estimation by Using LSTM in IMU-Based Gait Analysis—Proof of Concept. Sensors, 21(17), 5749. https://doi.org/10.3390/s21175749
23. Challa, S. K., Kumar, A., Semwal, V. B., & Dua, N. (2022). An Optimized-LSTM and RGB-D Sensor-Based Human Gait Trajectory Generator for Bipedal Robot Walking. IEEE Sensors Journal, 22(24), 24352–24363. https://doi.org/10.1109/jsen.2022.3222412
24. Rashmi, M., and Ram M., R., G., (2022). Human identification system using 3D skeleton-based gait features and LSTM model. Journal of Visual Communication and Image Representation, 82, 103416. https://doi.org/10.1016/j.jvcir.2021.103416
25. Semwal, V. B., Jain, R., Maheshwari, P., & Khatwani, S. (2023b). Gait reference trajectory generation at different walking speeds using LSTM and CNN. Multimedia Tools and Applications, 82(21), 33401–33419. https://doi.org/10.1007/s11042-023-14733-2
26. Guan, X., Chen, H., Liu, Y., Zhang, Z., & Ji, L. (2025). Predicting ground reaction forces and center of pressures from kinematic data in crutch gait based on LSTM. Medical Engineering & Physics, 139(1), 104338. https://doi.org/10.1016/j.medengphy.2025.104338
27. D'Souza, N., Charlton, J., Grayson, J., Kobayashi, S., Hutchison, L., Hunt, M., & Simic, M. (2022). Are biomechanics during gait associated with the structural disease onset and progression of lower limb osteoarthritis? A systematic review and meta-analysis. Osteoarthritis and Cartilage, 30(3), 381–394. https://doi.org/10.1016/j.joca.2021.10.010
28. Xu, D., Zhou, H., Quan, W., Jiang, X., Liang, M., Li, S., Ugbolue, U. C., Baker, J. S., Gusztav, F., Ma, X., Chen, L., & Gu, Y. (2024). A new method proposed for realizing human gait pattern recognition: Inspirations for the application of sports and clinical gait analysis. Gait & Posture, 107, 293–305. https://doi.org/10.1016/j.gaitpost.2023.10.019
29. Molavian, R., Fatahi, A., Abbasi, H., & Khezri, D. (2023). Artificial Intelligence Approach in Biomechanics of Gait and Sport: A Systematic Literature Review. Journal of Biomedical Physics and Engineering, 13(5), 383–402. https://doi.org/10.31661/jbpe.v0i0.2305-1621
30. Kannan, A. S., Hartwell, M. J., Grace, T., Hammond, E., Soriano, K. K. J., Souza, R. B., & Zhang, A. L.-Y. (2022). Correlating Biomechanical Gait Analysis with Patient-Reported Outcomes after Hip Arthroscopy for Femoroacetabular Impingement Syndrome. Orthopaedic Journal of Sports Medicine, 10(9), 23259671221121350. https://doi.org/10.1177/23259671221121352
31. Young-Shand, K. L., Roy, P. C., Dunbar, M. J., Abidi, S. S. R., & Astephen Wilson, J. L. (2023). Gait biomechanics phenotypes among total knee arthroplasty candidates by machine learning cluster analysis. Journal of Orthopaedic Research, 41(2), 335–344. https://doi.org/10.1002/jor.25363
32. Makihara, Y., Mannami, H., Tsuji, A., Hossain, Md. A., Sugiura, K., Mori, A., & Yagi, Y. (2012). The OU-ISIR Gait Database Comprising the Treadmill Dataset. IPSJ Transactions on Computer Vision and Applications, 4, 53–62. https://doi.org/10.2197/ipsjtcva.4.53