Robotica (2025), **43**, pp. 4174–4189
doi:10.1017/S0263574725102786

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Prediction of walking direction based on foot rotation angle features

Dianchun Bai[1,2], Jinli Dai[1], Zhuo Liu[1], Ruiji Li[1], Fenghao Gong[1], Zhongpeng Lin[1], Yinlai Jiang[2] and Hiroshi Yokoi[2]

[1]School of Electrical Engineering, Shenyang University of Technology, Shenyang, Liaoning, 110870, China
[2]Center for Neuroscience and Biomedical Engineering, University of Electro-Communications, Tokyo, 182-8585, Japan
Corresponding author: Dianchun Bai; Email: baidianchun@sut.edu.cn

Received: 4 March 2025; **Revised:** 24 September 2025; **Accepted:** 1 October 2025; **First published online:** 30 October 2025

Keywords: electronic walker; machine vision; depth learning; behavior prediction

Abstract

In the context of an aging population and declining birth rates, the advantages of robotic-assisted training are becoming increasingly prominent. However, improving the adaptability and safety of assistive walking robots remains a critical challenge. Accurately identifying a user's turning intent is essential for preventing dangerous situations such as falls or slips. As one of the core parameters of lower limb motion, foot rotation angles not only reflect the stability and coordination of gait but are also crucial for accurately predicting walking intentions, such as straight walking and turning. This study proposes a gated recurrent unit-based model for predicting foot rotation angles, driven by 3D visual data. By constructing a lower limb linkage model that includes foot joints and incorporating 3D foot rotation angle features, we develop a real-time algorithm for gait state prediction. This model enables accurate prediction of walking intentions, such as straight walking or turning, during walking and is experimentally validated using a robotic walker. The experimental results demonstrate the effectiveness of the proposed predictive model.

1. Introduction

Current research and development of walkers mainly focus on enhancing adaptive capabilities, such as automatic speed adjustment and obstacle avoidance functions [1]. Most existing intelligent walkers utilize contact-based sensors, such as pressure sensors, to capture human motion signals. Rokhmanova *et al*. [2] proposed a method for predicting foot position using knee adduction torque, aiming to alleviate knee osteoarthritis pain by improving the forward angle of the foot. Faraji *et al*. [3] introduced a new bipedal motion mechanical model that generates various walking gaits by simultaneously swinging the hip joint and supporting the ankle joint, driving the user's walking motion. However, these studies have overlooked the comprehensiveness and accuracy of real-time gait prediction, leading to situations where unexpected incidents cannot be effectively prevented.

The accurate recognition of human walking intentions is crucial for optimizing and enhancing the functionality of intelligent walkers. This not only helps improve walking efficiency but also enhances user experience and safety. To predict abnormal gait patterns, Brederecke *et al*. [4] analyzed gait freezing from accelerometer data worn by Parkinson's patients using a simple convolutional neural network, generating human-like lower limb trajectories under both periodic and non-periodic walking conditions. In particular, time-series neural networks, such as Long Short-Term Memory (LSTM) [5], have shown superior performance in processing gait time-series data. Kidziński *et al*. [6] used a wearable motion capture system to obtain data on foot heel strike and toe-off events during walking, predicting normal and pathological gaits in children by analyzing kinematic and labeled time-series data. An LSTM-based gait prediction model was developed by analyzing the results of gait cycle measurements. So, adjusting

© The Author(s), 2025. Published by Cambridge University Press.

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

the device based on the user's force like a customized stiffness control strategy is safer and more personalized [7]. Tie Liu *et al*. [8]. built a model of overall lower limb posture features during normal walking and applied it to lower limb motion posture prediction. However, these gait prediction methods do not address the prediction of user walking intentions. The introduction of gait prediction technology enables the walker to make adjustments based on predicted user behavior, greatly enhancing the walker's proactivity and foresight, thereby providing a safer walking experience for the user. However, these methods require data fusion from multiple sensors, making the data analysis complex and the user experience less comfortable.

In recent years, significant progress has been made in gait prediction algorithms based on 3D visual sensors. Malin-Mayor *et al*. [9] captured frontal video footage of individuals walking toward the camera and used AlphaPose [10] to estimate two-dimensional pose sequences. These time-series data were then input into a gait feature prediction neural network, enabling gait prediction and allowing clinicians and caregivers to intervene early to prevent falls. Molinaro *et al*. [11] used inertial measurement units and motion capture systems to collect and analyze gait data. Zang *et al.* [12] proposed a motion prediction network trained on joint node information. The dataset used, the Human 3.6M dataset, contains posture information such as joint angles. This dataset is derived from pose estimation applied to video recordings of human motion capture, providing full-body key joint angle data for a variety of movements. However, these methods require the acquisition of full-body node data, which results in large datasets and complex processing. A better way is to focus on local joints like lower limbs instead of the whole body – this simplifies control while keeping movement accuracy, as shown in humanoid robot motion research [13].

Focusing on the micro-actions during the walking process reveals new breakthroughs [14, 15]. Subtle changes in foot rotation and stride length at the moment of foot contact are closely related to gait intentions. By leveraging this relationship, gait analysis can be performed more accurately. These localized movements can serve as key features for distinguishing between different phases of gait.

Due to the relatively stationary position of the camera in relation to the user when using a walker, traditional gait features such as step length and stride are less pronounced, and the user's walking rhythm tends to be more consistent. This similarity in gait features reduces the effectiveness of distinguishing walking intentions such as turning or acceleration. As a result, it becomes difficult to differentiate between steps through simple pose comparisons, making it challenging to accurately identify the phase of each step within the overall gait cycle. Although significant progress has been made in gait analysis in other contexts, there remains a gap in research concerning gait feature differentiation and step phase recognition in walker usage scenarios, where the camera's stationary position relative to the user complicates these tasks.

In summary, current intelligent walkers still face numerous challenges in real-time walking intention prediction, data processing simplification, and improving user comfort. To reduce data processing load, enhance real-time performance, and improve comfort, while enabling adaptive adjustment capabilities in the walker, this study proposes a novel walking direction prediction algorithm based on foot rotation angle features, which focuses solely on lower limb joint nodes. This approach aims to reduce the user's operational burden, allowing them to focus more on their own walking.

To address the current issue that walkers cannot clearly guide the user's walking direction, this study proposes using a depth camera to capture lower limb point cloud data and estimate joint coordinates. The walking direction is determined based on the orientation of the ankle joint and the toe. To overcome the problem of not being able to directly capture continuous gait trajectories due to the relative motion between the camera and the user, this study introduces a method that captures the instantaneous turning angle of the toe. The angle's magnitude is then used to determine the user's forward direction. Finally, the constructed joint-linkage motion features are input into a gated recurrent unit (GRU) network to more effectively extract both the spatial features of the foot during walking and the temporal features between frames.

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

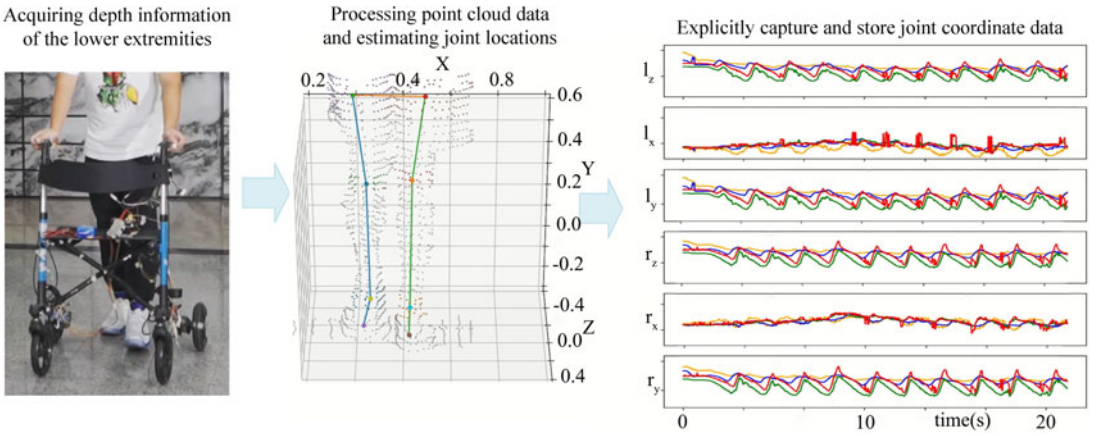


Figure 1. *Flowchart for obtaining lower limb joint nodes.*



Figure 2. *Walker hardware platform.*

2. Lower limb linkage model development

This paper proposes a 3D-vision-based lower limb linkage model, aimed at simplifying the complexity of walker operation and enhancing user comfort. Figure 1 illustrates the workflow for constructing the lower limb linkage model. First, the Intel RealSense D435i portable RGB-D camera is used to capture lower limb depth data, which is then converted into point cloud data. The point cloud data is subsequently optimized using multi-layer filters. To enhance the algorithm's processing speed, the panoramic point cloud is simplified through filters, including pass-through filtering, downsampling, and statistical filtering. The lower limb joint nodes, including the hip, knee, and ankle joint heights, are estimated to construct the lower limb linkage model [16]. In this study, the toe position is defined as the point closest to the camera below the ankle joint.

The walker experimental platform for collecting lower limb point cloud data utilizes a low-cost, lightweight walker developed by the research team, as shown in Figure 2. The walker is composed of a telescopic frame, a pair of hub-type permanent magnet synchronous electromagnetic motors, omni-directional wheels, a main controller, a suspended storage rack, a 3D visual sensor, and a perceptual handle. The visual sensor is positioned centrally below the suspended storage rack. The walker employs a vector control strategy on a single motor, achieving dual closed-loop control for current and speed. During straight-line motion, coordinated control is achieved using the same drive board. For turning,

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

a differential steering principle is applied to achieve coordinated control, while differential control and cross-coupling control principles are used to coordinate the two motors.

A linkage model is used to describe lower limb movement. To address the issue of joint displacement caused by static segmentation rules during walking, dynamic segmentation rules for limb joints are introduced. In static segmentation rules, joint positions are considered critical points between body segments with fixed spatial coordinates. However, in dynamic rules, the joint positions can experience vertical displacement (upward or downward) by y_{ofs}, where y_{ofs} denotes the vertical offset range ("ofs" is an abbreviation of "offset"). This parameter quantifies the maximum allowable deviation of joint positions from their reference locations in the static model, accommodating the physiological mobility of human joints during locomotion.

Accordingly, the cloud sets of the hip joint P_T, knee joint P_K, and ankle joint P_A, which describe the feasible ranges of vertical coordinates for each joint under dynamic conditions-are represented by Eq. (2a), (2b), and (2c).

$$P_T(x, y, z) = \{y \mid 1 - y_{ofs} \le y \le 1\} \tag{2a}$$

$$P_K(x, y, z) = \{y \mid 0.572 - y_{ofs} \le y \le 0.572 + y_{ofs}\} \tag{2b}$$

$$P_A(x, y, z) = \{y \mid 0.163 - y_{ofs} \le y \le 0.163 + y_{ofs}\} \tag{2c}$$

The spatial coordinates of the joints are obtained by solving for the centroid of the point cloud set of each lower limb node under the dynamic segmentation rules. The point with the maximum Z-axis value below the ankle is defined as the toe. Based on the proportion of each limb node in the static segmentation rules, given the thigh length l_{rh}, the shank length l_{rs}, and the inter-hip distance $2l_h$, the transformation matrices between lower limb joints in homogeneous coordinate form are derived as shown in Eq. (2d), (2e), and (2f).

The basic translation transformation matrix ${}^0_1\mathbf{T}$ (Eq. (2d)) serves as the foundation for spatial positioning in the lower limb linkage model, enabling accurate mapping from the global reference frame to the local hip coordinate system via homogeneous coordinates. Subscripts "0" and "1" denote coordinate systems: the 0 frame refers to the global reference frame fixed to the ground (with its origin set as the projection of the pelvic midline onto the plantar plane), and the 1 frame is the local coordinate system centered at the unilateral hip joint. The parameter l_h in the matrix represents the horizontal distance from the unilateral hip center to the midline of the global reference frame (h denotes hip), from which the bilateral hip distance is derived as $2l_h$. The fourth column element $[1, 0, 0, l_h]^T$ drives the coordinate system to translate along the x-axis by l_h, while the remaining rows ensure no translation in the y and z axes and maintain the rigidity of the coordinate system. The fourth row $[0, 0, 0, 1]$ is a signature feature of rigid-body transformations in homogeneous coordinates. This transformation decouples global motion from local joint motion; establishing the local hip coordinate system provides a unified coordinate framework for pose analysis of the knee and ankle joints, avoiding computational complexities caused by multi-joint coupling in the global reference frame.

$${}^0_1\mathbf{T} = {}^0_1\mathbf{P}(l_h, 0, 0) = \begin{bmatrix} 1 & 0 & 0 & l_h \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2d}$$

Transformation matrices of joints such as the knee and ankle follow the composite logic of translation describing limb length and rotation characterizing joint posture (as shown in Eq. (2e) and (2f)).

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

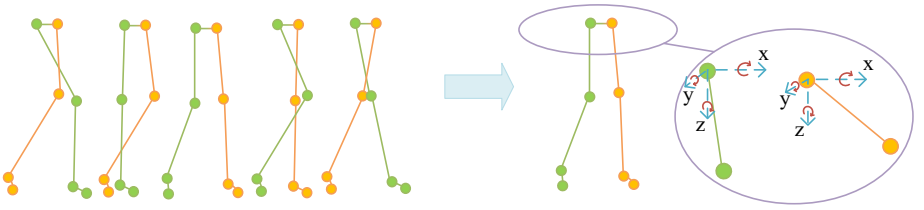


Figure 3. *Gait cycle and link rotational axes in the joint-linkage diagram.*

Taking the knee joint transformation matrix ${}^{1}_{2}\mathbf{T}$ (Eq. (2e)) as an example, subscripts "1" and "2" denote the hip coordinate system and knee coordinate system, respectively. As illustrated in the gait cycle schematic (Figure 3), the rotational part consists of the rotation matrix ${}^{1}_{2}\mathbf{R}(x, \frac{\pi}{2} - q_{zru})$ around the x-axis and the rotation matrix ${}^{1}_{2}\mathbf{R}(z, \frac{\pi}{2} - q_{xru})$ around the z-axis. The diagram clearly marks the x-axis (corresponding to the abduction-adduction direction) and z-axis (corresponding to the internal-external rotation direction) of the right thigh coordinate system, where q_{zru} and q_{xru} are joint attitude angles (q denotes angle; x and z indicate rotation axes; ru refers to right upper, i.e., the right thigh). They describe the abduction-adduction movement of the right thigh around the x-axis and the internal-external rotation around the z-axis, respectively. Here, the abbreviations "c" and "s" in the matrix represent cosine and sine functions. l_{rs} in the translation part ${}^{1}_{2}\mathbf{P}(0, l_{rs}, 0)$ represents the length of the shank (l denotes length; rs refers to the right shank), which translates the coordinate system along the y-axis by a corresponding distance, achieving the kinematic mapping of the rigid link (shank) + flexible joint (knee joint).

$$
\begin{aligned}
{}^{1}_{2}\mathbf{T} &= {}^{1}_{2}\mathbf{R}\left(x, \frac{\pi}{2} - q_{xru}\right)^{1}_{2}\mathbf{R}\left(z, \frac{\pi}{2} - q_{zru}\right)^{1}_{2}\mathbf{P}\,(0, l_{\mathrm{rh}}, 0) \\
&= \begin{bmatrix}
sq_{zru} & -cq_{zru} & 0 & -l_{\mathrm{rh}}cq_{zru} \\
cq_{zru}sq_{xru} & sq_{xru}sq_{zru} & -cq_{xru} & l_{\mathrm{rh}}sq_{xru}sq_{zru} \\
cq_{xru}cq_{zru} & cq_{xru}sq_{zru} & sq_{xru} & l_{\mathrm{rh}}cq_{xru}sq_{zru} \\
0 & 0 & 0 & 1
\end{bmatrix}
\end{aligned}
\tag{2e}
$$

The ankle joint transformation matrix ${}^{2}_{3}\mathbf{T}$ (Eq. (2f)) works similarly, with subscripts "2" and "3" denoting the knee coordinate system and ankle coordinate system, respectively. The combination of rotation matrices describes the posture change of the foot relative to the shank, and the translation matrix conveys information such as shank length. Through recursive multiplication of homogeneous transformations ${}^{0}_{3}\mathbf{T} = {}^{0}_{1}\mathbf{T} \cdot {}^{1}_{2}\mathbf{T} \cdot {}^{2}_{3}\mathbf{T}$, the complete pose relationship between the hip, knee, and ankle can be derived. This method, which converts complex physiological movements of the human lower limbs into quantifiable matrix operations, provides a solid mathematical foundation for applications such as gait cycle analysis and exoskeleton motion control.

$$
\begin{aligned}
{}^{2}_{3}\mathbf{T} &= {}^{2}_{3}\mathbf{R}\left(x, q_{zrd}\right)^{2}_{3}\mathbf{R}\left(z, \pi - q_{xrd}\right)^{2}_{3}\mathbf{P}\,(0, l_{rs}, 0) \\
&= \begin{bmatrix}
-cq_{zrd} & -sq_{zrd} & 0 & -l_{rs}sq_{zrd} \\
cq_{zrd}sq_{xrd} & -cq_{zrd}cq_{xrd} & -sq_{zrd} & -l_{rs}cq_{zrd}cq_{xrd} \\
sq_{zrd}sq_{xrd} & -cq_{xrd}sq_{zrd} & cq_{zrd} & -l_{rs}cq_{xrd}sq_{zrd} \\
0 & 0 & 0 & 1
\end{bmatrix}
\end{aligned}
\tag{2f}
$$

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

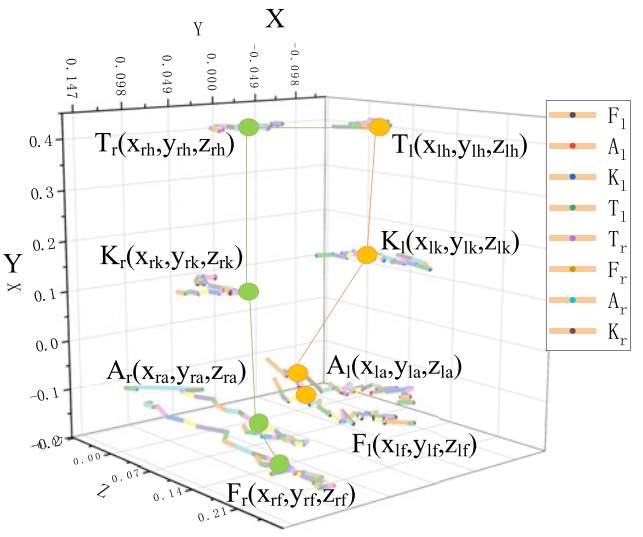


Figure 4. *Changes in walking motion trajectories of lower extremity joint points.*

Therefore, the spatial coordinates of each joint in the lower limb linkage model can be obtained as follows:

$$T_r = {}^{0}_{1}\mathbf{T}O\left(x_0, y_0, z_0\right) \tag{2g}$$

$$K_r = {}^{1}_{2}\mathbf{T}T_r(x_m, y_{rh}, z_{rh}) \tag{2h}$$

$$A_r = {}^{2}_{3}\mathbf{T}K_r\left(x_{rk}, y_{rk}, z_{rk}\right) \tag{2i}$$

The motion trajectory variations of the lower limb joints are shown in Figure 4, where T_r and T_l represent the right and left hip joints, K_r and K_l represent the right and left knee joints, A_r and A_l represent the right and left ankle joints, and F_r and F_l represent the right and left toes.

2.1. Lower limb linkage model gait feature extraction

Using an example of a walking sequence with straight paths and turns, Figure 5 shows gait cycle point cloud data and estimated joint data from lower limb linkages. Figure 5(a) shows the movement trajectories of lower limb joints during straight walking and turning, respectively. Figure 5(b) illustrates the changes in lower limb point cloud data and joint data during straight walking and turning.

To further elaborate on the data features utilized in this study, emphasis is placed on the toe coordinates and the steering Euler angles of the linkage between the toe and the ankle joint. These Euler angles precisely depict the rotational characteristics of foot movement, playing a pivotal role in comprehending the dynamic alterations in gait and posture. Figure 6 illustrates the timing sequence of left and right feet leaving the ground and landing within a gait cycle, along with the three-dimensional visualization of foot steering. This diagram offers an intuitive grasp of the rotational motion of the toe-ankle linkages across different gait phases. The actual foot photos and the views from the side, top, and front, respectively, showcase the changes in foot rotation during the gait cycle. In these views, α, β, and γ denote the angular variations of the toe-ankle linkage rotation around the X, Y, and Z coordinate axes, corresponding to the foot rotation angles.

Pitch (X-axis): Upward or downward motion of the toes (minimal variation during walking).

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

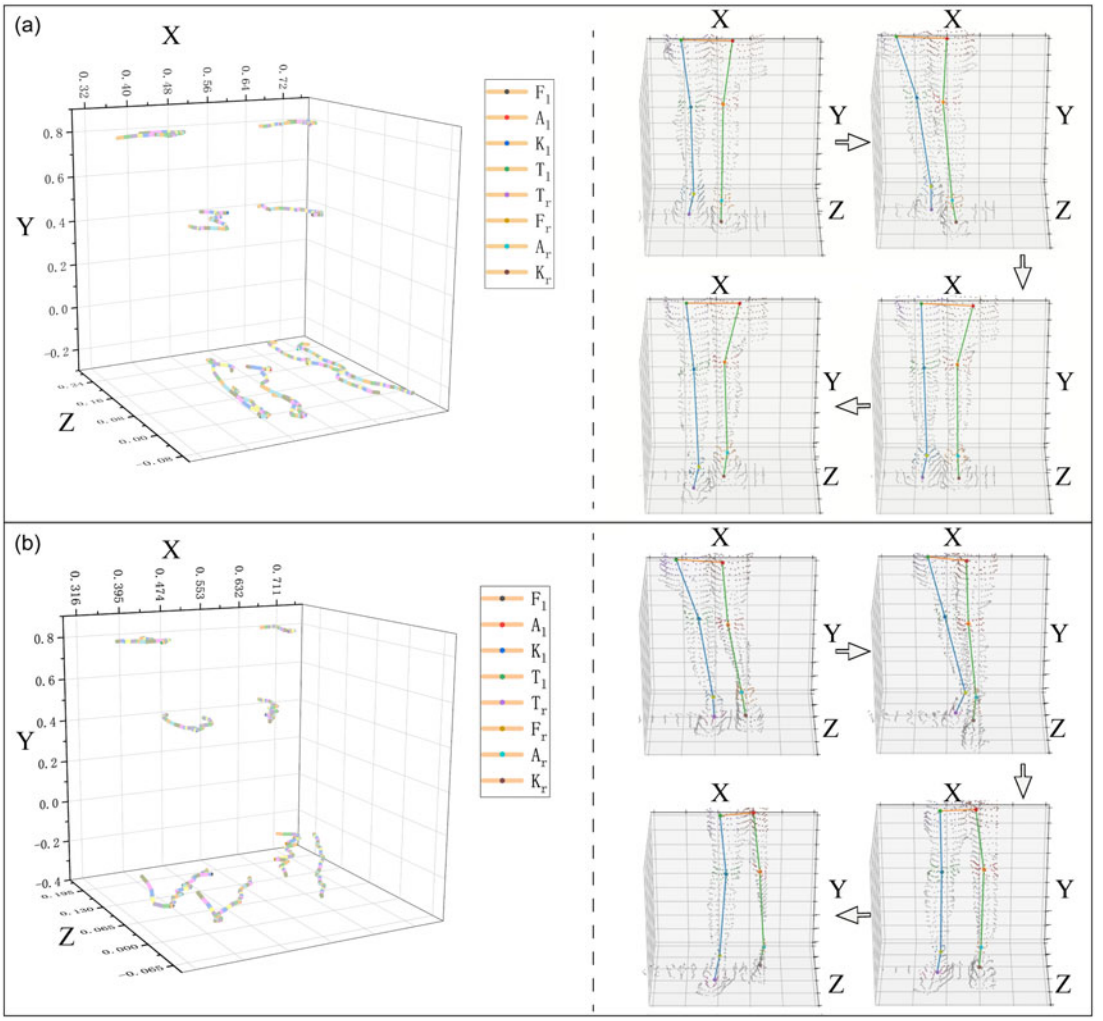


Figure 5. *Lower limb point cloud and joint linkages during straight walking and turning gait cycles: (a) straight; (b) left turn.*

Yaw (Y-axis): Left or right rotation of the toes (significant variation during turning).

Roll (Z-axis): Inward or outward tilt of the foot (moderate variation during walking).

Figure 7 presents the dynamic variations in the toe coordinates and foot rotation angles during the walking cycle. Specifically, panels (a) and (b) depict the 3D coordinates of the left and right toes, respectively, capturing the positional changes of the toe joints along the X, Y, and Z axes throughout the gait. Panels (c) and (d) illustrate the corresponding Euler angles of foot rotation, which reflect the rotational behavior of the left and right foot relative to the three primary coordinate axes (X, Y, and Z). These angular changes are crucial in understanding the kinematic properties of foot movement during walking.

The variation in the 3D coordinates of the toe is shown in Figure 7(c) and 7(d). Within 10–15 s, which correspond to the left-turning phase, a noticeable change in the toe coordinates can be observed during the turning motion. After 20 s, the stacked data reveals a periodic fluctuation in the foot data, indicating that focusing on the foot's gait features is effective. During walking, in addition to the periodic change in joint positions, the angular rotation of the linkages exhibits a more pronounced periodic oscillation. The variation in the 3D Euler angles of foot rotation is shown in Figure 7(c) and 7(d), where the amplitude

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

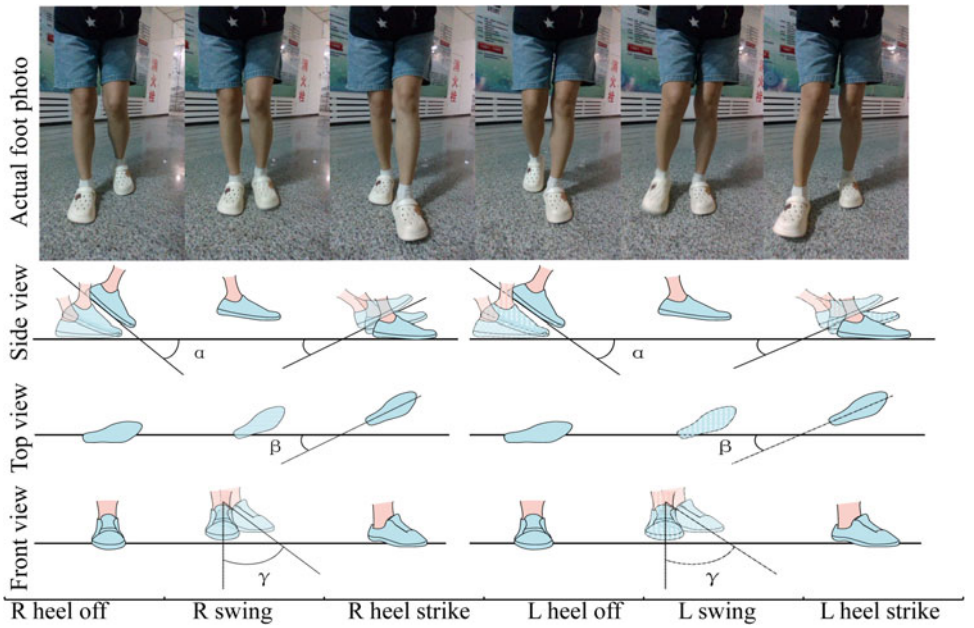


Figure 6. *Gait cycle foot motion: liftoff-landing and 3D steering.*

of the angle changes becomes especially significant during the turning phase. This angular variation can serve as a key feature for determining the occurrence of turns.

2.2. Walking intention prediction model development

A prediction model for the lower limb center of mass (COM) and steering posture angles based on GRU is proposed. The COM and steering posture angle data during the lower limb movement exhibit time-series characteristics, making the GRU model suitable for prediction. The GRU model, with its reset and update gates, addresses the gradient vanishing problem in traditional RNNs and offers faster training speeds. The network structure is shown in Figure 8. The GRU incorporates the concepts of the Reset Gate and Update Gate, modifying the calculation of hidden states in recurrent neural networks.

Given that different parts of the gait data, when the user pushes the walker, have varying levels of importance for prediction, an attention mechanism is employed to automatically learn and focus on the most critical information for current gait predictions. For instance, during a turning motion, the focus is placed on the ankle and toe regions. At each time step, attention weights are calculated for the input features, highlighting significant joint angle changes, which improves the model's ability to handle ambiguous gait features and enhances prediction accuracy.

As highlighted in Figure 8, the yaw angle emerges as the primary feature for detecting the direction of turning movements, consistent with observations from the gait analysis. Our model is specifically designed to take the yaw angle-related features as the core, while incorporating 6-dimensional data of 3D rotational Euler angles (including pitch, yaw, and roll) from both feet. This design choice was made after comparative evaluations: while a stand-alone yaw angle can capture the primary turning intent, the inclusion of pitch and roll angles provides supplementary information about subtle foot adjustments during complex gaits (e.g., body tilt compensation in sharp turns). Such a multidimensional feature set, without introducing other leg keypoints, aims to fully exploit the most representative foot motion information while enhancing prediction robustness.

x_t: Input data at the current time step.

h_{t-1}: Hidden state from the previous time step.

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

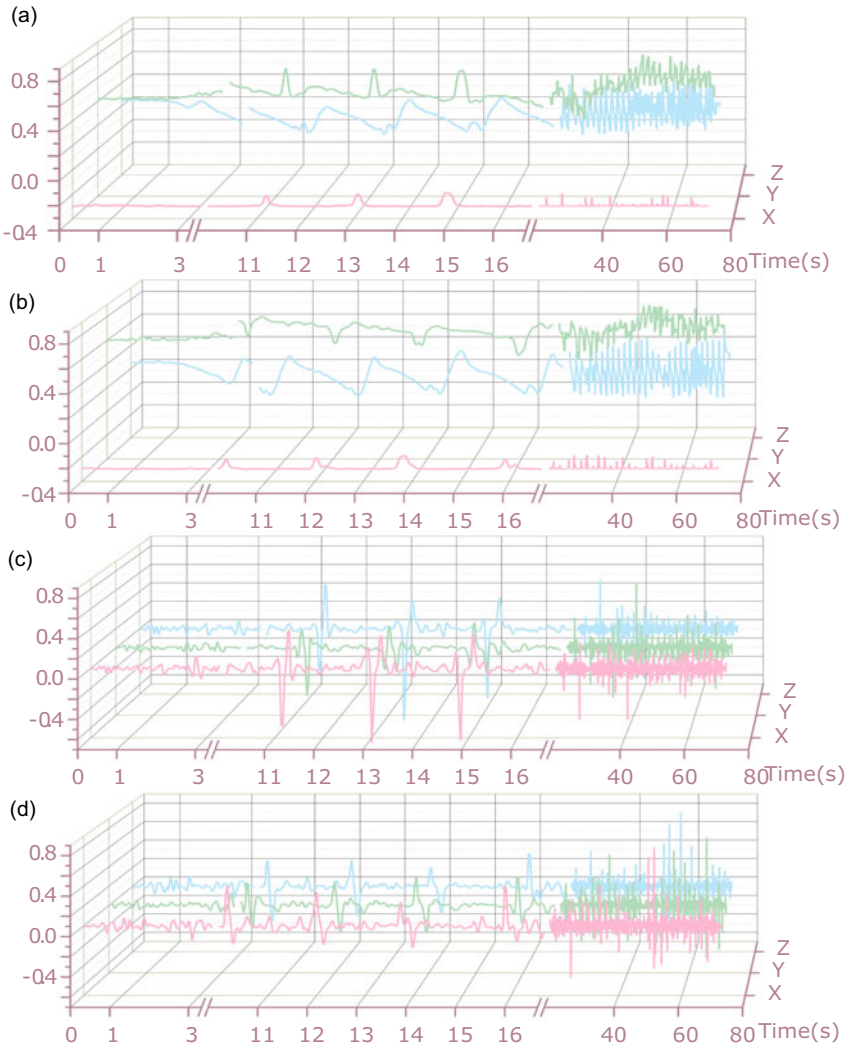


Figure 7. *Changes in toe coordinates and foot rotation Euler angles during walking: (a) 3D coordinates of the left toe; (b) 3D coordinates of the right toe; (c) Euler angles of left foot rotation; (d) Euler angles of right foot rotation. (Caption: The pink, green, and blue colors represent the X, Y, and Z axis coordinates of the toe joint, while the pink, green, and blue colors represent the Euler angles of rotation around the X, Y, and Z axes).*

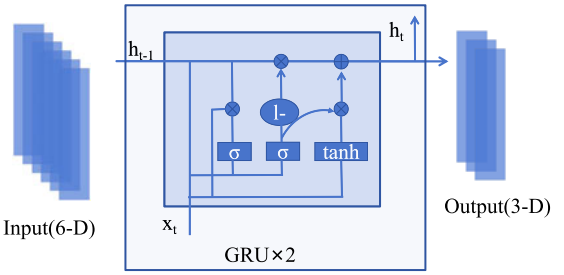


Figure 8. *Architecture of the GRU network.*

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

Table I. *Comparison of training parameters across all prediction models.*

Parameter	GRU	LSTM	BiGRU	BiLSTM	Attention-GRU	Transformer
Optimizer	Adam	Adam	Adam	Adam	Adam	Adam
Learning rate	0.001	0.001	0.001	0.001	0.001	0.001
Weight decay	1e-5	1e-5	1e-5	1e-5	1e-5	1e-5
Hidden layers	2	2	2	2	2	-
Encoder layers	-	-	-	-	-	2
Hidden size	64	64	64	64	64	64
Bidirectional	No	No	Yes	Yes	No	-
Attention heads	-	-	-	-	-	4
Attention mechanism	No	No	No	No	Yes	Yes
Training epochs	100	100	100	100	100	100

r_t, u_t: Current output, calculated as shown in Eq. (3a) and (3b) (values range from 0 to 1).

$$r_t = \sigma\ (W_r\ [h_{t-1}, x_t] + b_r) \tag{3a}$$

$$u_t = \sigma\ (W_u\ [h_{t-1}, x_t] + b_u) \tag{3b}$$

The candidate hidden state is given by Eq. (3c), where it is controlled by the reset gate from the previous time step's historical information. This allows for the selective retention or discarding of all historical information up to the previous time step.

$$h'_t = \tanh\ (W_c\ [r_t \odot h_{t-1}, x_t] + b_c) \tag{3c}$$

The current hidden state, as shown in Eq. (3d), resolves the gradient vanishing problem in RNNs.

$$h_t = (1 - u_t) \odot \tanh\left(h'_t\right) + u_t \odot h_{t-1} \tag{3d}$$

The model combines the GRU with a single-layer fully connected neural network to map the hidden layer features. For fair comparison, key hyperparameters (optimizer, learning rate, training epochs, etc.) were standardized across all evaluated models. Table I details the architectural and training parameters for GRU, LSTM, BiGRU, BiLSTM, Attention-GRU, and Transformer models – selected to contrast recurrent versus non-recurrent designs, bidirectional versus unidirectional processing, and attention mechanism effects in gait intent prediction.

3. Experiments

3.1. Dataset creation

To ensure the representativeness of the dataset and the generalization capability of the model, this study involved 15 healthy university students (10 males and 5 females, aged 20–25) as participants for lower limb motion data collection. An Intel RealSense D435i vision sensor was installed on the walker directly below the central axis of the suspended storage rack, positioned vertically to the ground, and facing the user at a distance of approximately 50 cm and a height of about 37 cm from the ground. The lower limb motion data collection system operated on the Windows 11 operating system, with code development in Visual Studio 2019 and PyCharm. The software utilized the Intel RealSense SDK 2.0, OpenCV 4.4.5, and PCL 1.10.1 libraries. The hardware environment included an Intel(R) Core(TM) i5-13500H CPU @ 2.60 GHz and an NVIDIA GeForce RTX 4060 GPU.

After selecting the gait angles, volunteers were asked to push the walker forward indoors under conditions of adequate lighting and a flat surface, with the assistive robot set to a constant-speed forward mode. Each experiment involved pushing the walker for a distance of 2 m. Measurements showed that

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

Table II. *Model performance comparison.*

Model	MAE	MSE	Size (KB)	Training time (s)	Test time (s)
GRU	4.14×10^{-2}	5.12×10^{-3}	0.52×10^{5}	35.21	0.04
LSTM	3.82×10^{-2}	4.85×10^{-3}	0.69×10^{5}	38.21	0.06
GRU_Attention	5.01×10^{-2}	8.35×10^{-3}	0.52×10^{5}	56.10	0.03
BiGRU	4.11×10^{-2}	5.18×10^{-3}	1.28×10^{5}	44.69	0.06
BiLSTM	3.72×10^{-2}	4.50×10^{-3}	1.70×10^{5}	74.67	0.05
Transformer	5.05×10^{-2}	8.53×10^{-3}	5.67×10^{5}	267.43	0.11

during slow forward movement, one gait cycle lasted approximately 2.5 s. Additionally, depth images were collected at a rate of 30 frames per gait cycle to capture the full walking cycle. Since there were three walking directions, and each direction was repeated 10 times, a total of 2250 depth images were collected for each participant ($30 \times 2.5 \times 3 \times 10$). For 15 participants, a total of 33,750 frames of depth data were collected, with each frame estimating the 3D data of 8 joint positions and the 3D rotational angles of 6 linkages connecting the joints. This resulted in the collection of 1,417,500 data points in total. The resulting samples were then divided into a training set, a validation set, and a test set with a ratio of 7:1:2.

The data is normalized to improve the convergence speed and accuracy of the model. To preserve the integrity of the temporal features, a sliding window approach is used for data sampling. Every m consecutive frames are treated as one input sample, with a prediction step of n. The (m + n)-th frame serves as the label for that sample. The sliding window starts from the first frame of the sample data and moves until the label reaches the last frame of the data.

3.2. Prediction experiment and result analysis

In this experiment, three evaluation metrics are used to describe the error between the test sample data and the model's output: mean absolute error (MAE), mean squared error (MSE), and root MSE. To comprehensively validate the predictive capability of the GRU network, we expanded the comparison to include more relevant temporal sequence models, including LSTM, GRU_Attention, BiGRU, BiLSTM, and Transformer. These models were selected for their suitability in temporal gait data modeling: LSTM and GRU as classical recurrent neural networks with mature performance in capturing temporal dependencies; bidirectional structures (BiGRU, BiLSTM) for capturing both forward and backward temporal information aligned with walking continuity; GRU_Attention to enhance focus on key temporal nodes for gait transitions; and Transformer for its advantages in long-term temporal dependency modeling and recent applications in human motion analysis. All compared models share the same number of layers to ensure fair evaluation of backpropagation-based prediction accuracy.

As shown in TABLE II, in the test set samples, the GRU network maintains a favorable balance between predictive performance and resource efficiency. While BiLSTM achieves the smallest MAE and MSE, and LSTM also performs well, GRU exhibits comparable error metrics with a smaller size and shorter training time. GRU_Attention and Transformer show larger errors, with Transformer requiring significantly more resources. BiGRU, though comparable in error to GRU, has a larger model size. Given the limited computational resources of the control module, the GRU model is chosen for its optimal trade-off between prediction accuracy, model scale, and efficiency. Figure 9 presents the training loss curves of all comparative models. As epochs increase, the GRU model exhibits faster convergence to a lower stable loss, greater training efficiency than attention-augmented models (Attention-GRU, Transformer). This indicates the GRU's advantages in balancing computational simplicity and predictive accuracy for gait sequence data, validating its suitability as the core architecture in this study.

Taking the prediction of the left foot rotation angle during a turning cycle as an example (Figure 10), the visualization of test set results shows that the GRU model's predicted curve aligns closely with the

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

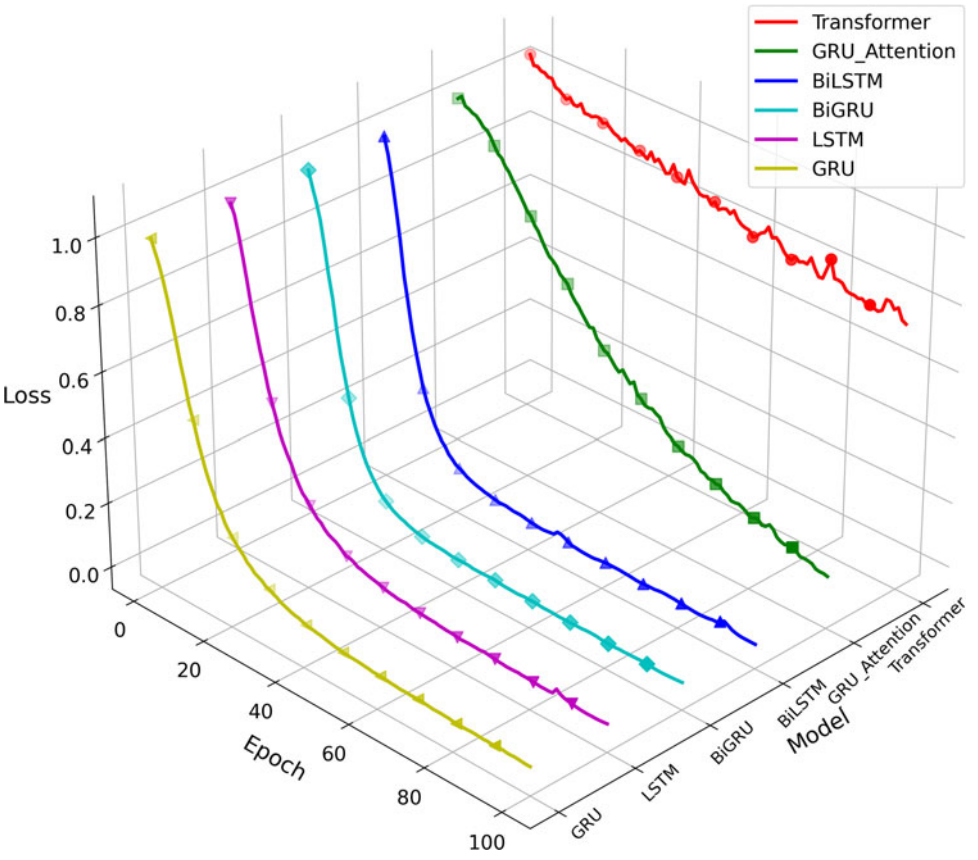


Figure 9. *Loss value changes during model training.*

actual data trend – consistently capturing key variations and even anticipating trend changes 6 frames in advance. This not only confirms the model's strong predictive performance in tracking overall data trends but also highlights the GRU's effectiveness in handling gait-related sequence prediction tasks compared to other evaluated architectures.

The predicted and actual foot rotation angles throughout the entire walking route are presented in Figure 11, where Figure 11(a) and 11(b) corresponds to the left and right feet, respectively. Based on the experimental results, a turn is detected when the foot rotation angle exceeds 54° (equivalent to $0.3 \times 180°$ after normalization, with 180° and −180° scaled to 1 and −1, respectively). The turning detection logic is designed to be robust: a turn is triggered if the rotation angle of either the left or the right foot exceeds the threshold. The turning direction (left or right) is determined by the rotation variations of the left and right feet. For instance, as observed between 12 and 18 s in Figure 11, although the prediction for the right foot is less accurate, the left foot's rotation is reliably predicted and consistently exceeds the threshold, enabling the system to correctly identify a left turn. As indicated by the predicted trends and actual angles, the assistive walking device gradually adjusts the turning angle, thereby enabling effective assistance for both walking and turning functions.

To verify the accuracy of the proposed method, the dynamic time warping (DTW) algorithm is used to assess the prediction model's accuracy. As Yadav and Alam [17] noted, DTW can nonlinearly align time series, effectively measuring action sequences with different lengths but similar patterns, and can robustly capture the overall trend of action sequences rather than matching exact time points. DTW computes the distance between two time series, where the distance serves as a quantification of their similarity. A smaller distance value indicates a higher degree of matching between the two time series after time alignment, meaning they are more similar.

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

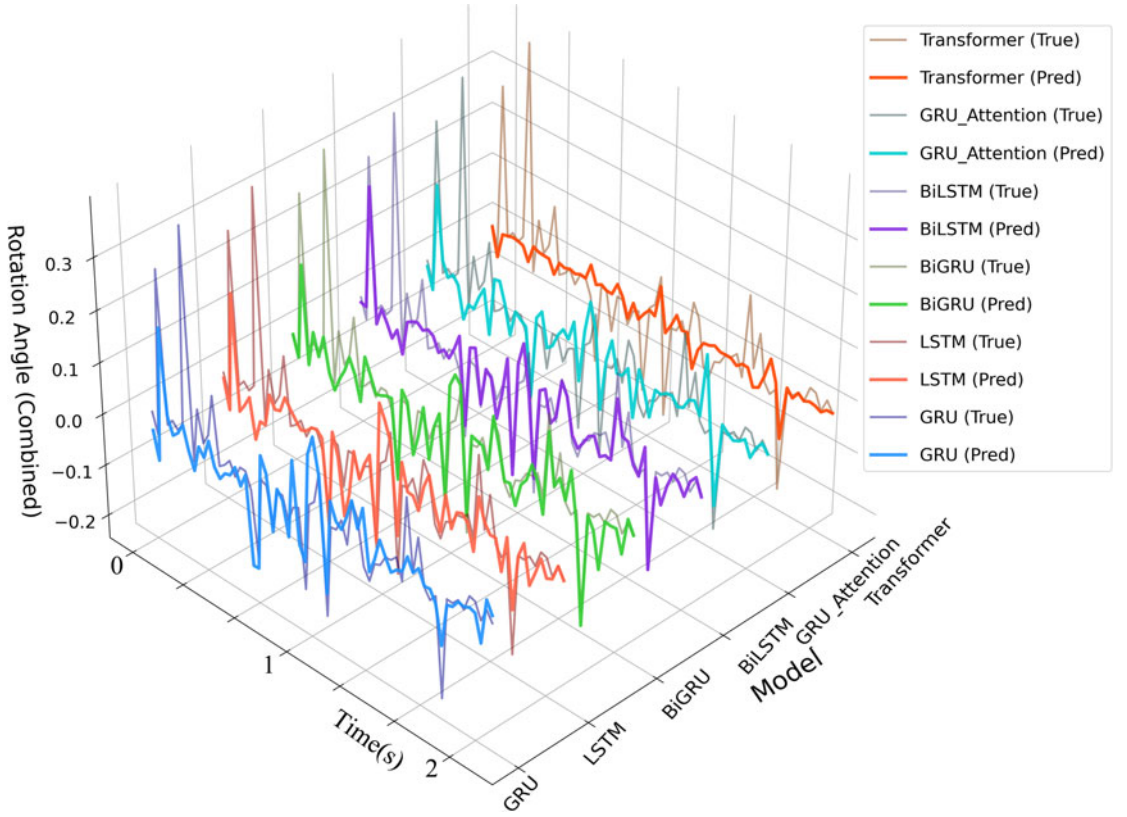


Figure 10. *Prediction of the angle of rotation of the left foot during a steering cycle.*

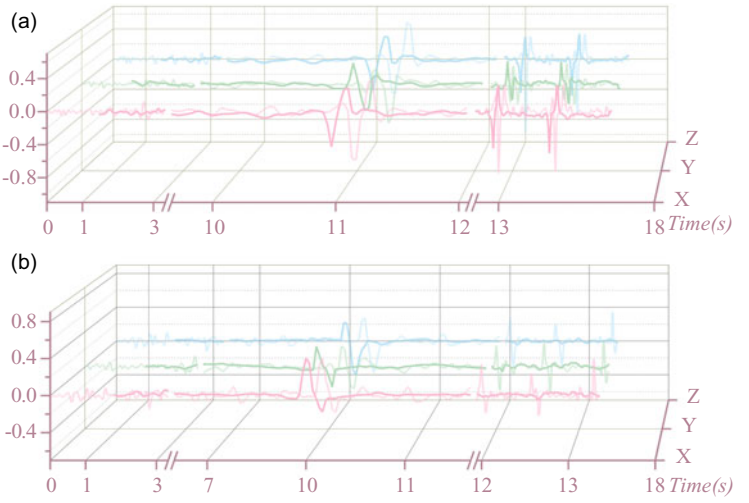


Figure 11. *Foot rotation angles: actual (transparent) and predicted (solid). (a) left; (b) right.*

The windowed data of the toe and ankle joint linkage during walking are set as $A = [x_1, x_2, x_3, \ldots, x_m]$ and predicted time-axis data windowed as $B = [y_1, y_2, y_3, \ldots, y_m]$, where (x, y) represents any point in the matrix. Directly calculating the distance between sequences A and B may be inaccurate due to potential temporal shifts (phase differences) between them. Therefore, actual and predicted data were aligned first, as illustrated in Figure 12. Figure 12(a) and (b) shows the original unaligned real and predicted

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

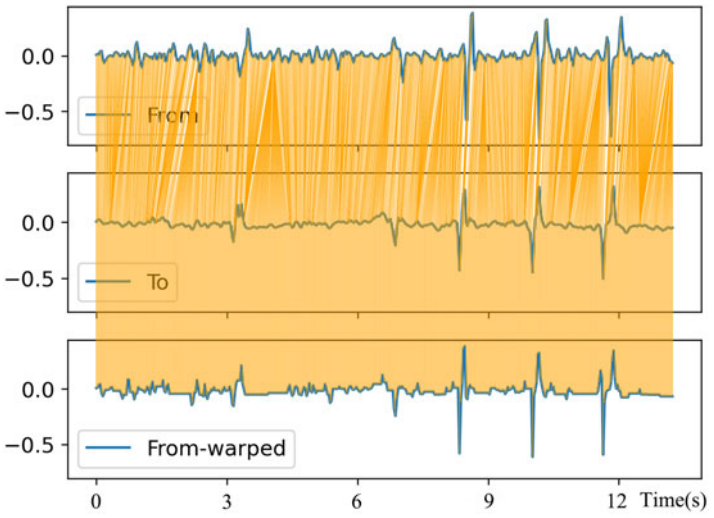


Figure 12. *Data alignment process illustration: (a) real data; (b) predicted data; (c) aligned data.*

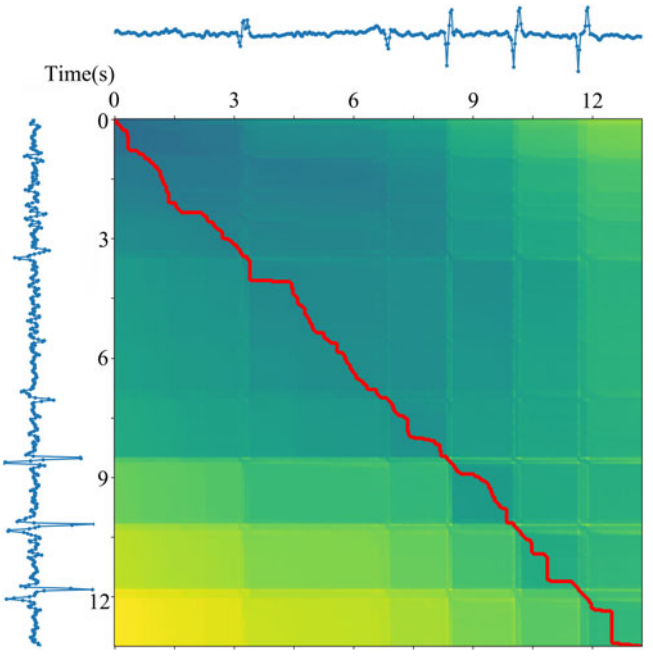


Figure 13. *DTW validates predictive model accuracy.*

sequences, respectively, highlighting their temporal misalignment. The alignment process finds an optimal warping path to match corresponding points between the two sequences, resulting in the aligned data shown in Figure 12(c). The DTW distance was calculated as *Distance* = 0.9848, as shown in Figure 13. This demonstrates a high level of similarity between the time series, verifying the effectiveness of the proposed algorithm.

3.3. Conclusion

This study focuses on the key technical challenges faced by lower limb rehabilitation robots in assisting individuals with lower limb disabilities during walking. For the first time, we propose a user walking

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

intention prediction method based on foot rotation angles. Utilizing point cloud depth vision technology, we accurately capture rich and detailed lower limb motion data, which provides a robust data foundation for subsequent model development. By incorporating multi-path GRU, we construct a foot rotation angle prediction model and establish a lower limb posture feature model for users during typical walking states such as straight walking and turning. Experimental results validate the proposed model's advantages, including its low resource demand and high accuracy, demonstrating its ability to predict walking direction precisely in complex real-world scenarios.

However, improvements are still needed: future work will expand the participant cohort to include older adults and individuals with disabilities, systematically testing the generalizability of our findings to enhance practical applicability. Additionally, we will quantify inference time, resource consumption on embedded systems for assistive devices, and prioritize optimizing real-time performance – addressing these aspects to strengthen the model's translational value in clinical and daily use settings.

Author contributions. D.B., the primary designer of the research framework, provided the experimental approach and guidance on the manuscript. He was responsible for the overall direction and coordination of the project. J.D., a female contributor, was in charge of programming, conducting experiments, and drafting the manuscript. She played a central role in data collection and analysis, ensuring the accuracy and reliability of the experimental results. Z.L., R.L., Z.L., and F.G. collaborated on data collection, ensuring the integrity and quality of the dataset. They also provided a critical review of the manuscript, offering feedback on data presentation and interpretation. Y.J. and H.Y. offered valuable advice and guidance throughout the research and writing process. They helped shape the direction of the study with their professional insights and contributed to the refinement of the manuscript through rigorous review and suggestions. All authors have contributed to the research work and have read and approved the final version of the manuscript.

Financial support. This study was jointly supported by the Chunhui Program of the Ministry of Education 202200215.

Competing interests. The authors declare that no conflicts of interest exist.

Ethical approval. Not applicable.

References

[1] T. Linder, N. Vaskevicius, R. Schirmer and K. O. Arras, "Cross-Modal Analysis of Human Detection for Robotics: An Industrial Case Study," **In:** IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE, 2021) pp. 971–978.

[2] N. Rokhmanova, K. J. Kuchenbecker, P. B. Shull, R. Ferber and E. Halilaj, "Predicting knee adduction moment response to gait retraining with minimal clinical data," *PLOS Comput. Biol.* **18**(5), e1009500 (2022).

[3] S. Faraji and A. J. Ijspeert, "Scalable closed-form trajectories for periodic and non-periodic human-like walking," **In:** *International Conference on Robotics and Automation (ICRA)* (IEEE, 2019) pp. 5295–5301.

[4] J. Brederecke, "Freezing of gait prediction from accelerometer data using a simple 1D-convolutional neural network – 8th place solution for Kaggle's parkinson's freezing of gait prediction competition," arxiv preprint arxiv:2307.03475 (2023).

[5] B. Su and E. M. Gutierrez-Farewik, "Gait trajectory and gait phase prediction based on an LSTM network," *Sensors* **20**(24), 7127 (2020).

[6] Ł. Kidziński, S. Delp and M. Schwartz, "Automatic real-time gait event detection in children using deep neural networks," *PloS one* **14**(1), e0211466 (2019).

[7] A. Kapsalyamov, S. Hussain, R. Goecke, N. A. Brown and P. K. Jamwal, "Customized stiffness control strategy for a six-bar linkage-based gait rehabilitation robot," *Robotica* **42**(10), 3398–3415 (2024).

[8] T. Liu, D. Bai, H. Yi and H. Yokoi, "Prediction of abnormal gait behavior of lower limbs based on depth vision," *Robotica* **43**(1), 47–68 (2025).

[9] C. Malin-Mayor, V. Adeli, A. Sabo, S. Noritsyn, C. Gorodetsky, A. Fasano, A. Iaboni and B. Taati, "Pose2Gait: Extracting Gait Features From Monocular Video of Individuals with Dementia," **In:** *International Workshop On PRedictive Intelligence In MEdicine 2023*, (Springer, 2023) pp. 265–276.

[10] H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li and C. Lu, "Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time," *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(6), 7157–7173 (2023).

[11] D. D. Molinaro, I. Kang, J. Camargo, M. C. Gombolay and A. J. Young, "Subject-independent, biological hip moment estimation during multimodal overground ambulation using deep learning," *IEEE Trans. Med. Robot.* **4**(1), 219–229 (2022).

[12] C. Zang, M. Pei and Y. Kong, "Few-shot Human Motion Prediction Via Learning Novel Motion Dynamics," **In:** Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence 2021 (IJCAI, 2021) pp. 846–852.

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press

[13] S. Yu, Y. Nakata, Y. Nakamura and H. Ishiguro, "Joint angle synergy-based humanoid robot motion generation with fascia-inspired nonlinear constraints," *Robotica* **42**(8), 2566–2588 (2024).
[14] N. Mani, P. Haridoss and B. George, "A wearable ultrasonic-based ankle angle and toe clearance sensing system for gait analysis," *IEEE Sens. J.* **21**(6), 8593–8603 (2021).
[15] G. Yang, J. Heo and B. B. Kang, "Adaptive vision-based gait environment classification for soft ankle exoskeleton," *In Actuators*, **13**(11), 428 MDPI (2024).
[16] Standardization Technical Committee for Ergonomics of China (SAC/TC7), "Gb/t 10000-2023: Anthropometric dimensions for adults in China," (Beijing, Standards Press of China, 2023).
[17] M. Yadav and M. A. Alam, "Dynamic time warping (DTW) algorithm in speech: A review," *IJRECE* **6**(1), 524–528 (2018).

Cite this article: D. Bai, J. Dai, Z. Liu, R. Li, F. Gong, Z. Lin, Y. Jiang and H. Yokoi (2025). "Prediction of walking direction based on foot rotation angle features", Robotica **43**, 4174–4189. https://doi.org/10.1017/S0263574725102786

https://doi.org/10.1017/S0263574725102786 Published online by Cambridge University Press