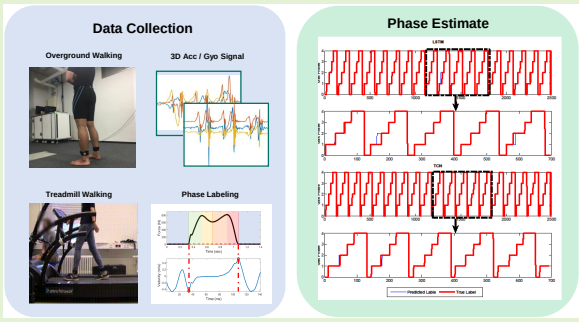
Speed-Variable Gait Phase Estimation During Ambulation via Temporal Convolutional Network

Yan Guo, Yonatan Hutabarat, Dai Owaki, *Member, IEEE,* and Mitsuhiro Hayashibe, *Senior Member, IEEE*

***Abstract*— Accurate estimation of gait phases during walking is a crucial prerequisite for both extracting clinically meaningful gait parameters and delivering gait-based feedback control information to rehabilitation devices. In addition, speed variation appears in our daily walking locomotion. However, most existing IMU-related methods based on heuristic algorithms were reported to be sensitive to walking speed changes. To address this problem, in this study, we presented a temporal convolutional network (TCN) based approach for automatic and robust recognition of gait phases across multi-scene ambulation and different walking speeds. We collected data on both real-world overground walking experiments and public treadmill datasets to validate the performance in terms of the accuracy and robustness of the proposed method. By comparing our method with six machine learning models and two neural network models, our method achieves 97% accuracy in gait phase estimation for both overground and treadmill walking, outperforming all compared benchmarks. It also excelled in both model generalizability evaluation and velocity robustness comparison tests over the other two neural networks. Notably, TCN can achieve 91 % accuracy in velocity robustness tests and outperformed fully convolutional network (FCN) and long short-term memory (LSTM) in mean square error comparison ($p<0.05$). These results show that our method has outstanding estimation performance and high robustness on gait speed variations.**



***Index Terms*— Gait Phase Estimation, Gait Speed Variation, Wearable Sensors, Multi-Scene Ambulation, Deep Learning**

I. Introduction

INDEPENDENT walking is one of the primary human physical activities among daily living activities completed by the coordination of multiple muscles and skeletons. The gait cycle describes the cyclic pattern of movement that occurs while walking, which is defined as a single cycle of gait that starts when the heel of one foot strikes the ground and ends when the same heel touches the ground again [1], [2]. In order to clearly quantify the gait cycle, researchers divided the gait process into multiple phases. The most studied model in the gait analysis field was a separation of the stride into the stance phase and the swing phase. Stance (ST), in which the foot is in contact with the ground and bearing body weight, and swing (SW) is the period foot is free to move forward. Detailed further, the stance phase comprises several sub-phases: initial contact or heel strike, loading response, midstance, terminal stance, and pre-swing [3].

The importance of gait phase segmentation lies in its ability to evaluate and diagnose gaits. Automatic gait phase segmentation is a challenging task that is of increasing interest in various multidisciplinary fields. For example, in the sports field, gait analysis can serve as an effective tool for coaching athletes to improve their sports performances and prevent injuries [4], [5]. Gait analysis is also indispensable to support clinical diagnosis, wearable exoskeleton control, and rehabilitation process monitoring. For instance, due to factors such as age and neurodegenerative diseases, spatiotemporal gait parameters can degenerate significantly from the typical patterns [6]. Events discriminating gait phases can also be used as control inputs for functional electrical stimulation (FES) and wearable assistive devices, such as prostheses, orthoses, and exoskeletons [7]. Monitoring the improvement of spatiotemporal gait parameters can also be considered an effective method of rehabilitation.

However, current gait phase analysis relies on subjective clinician inspection, which requires field-specific expertise to hand-craft features for phase extraction. This approach is not only tedious and time-consuming but also increases the potential for human error [8]. Therefore, there is a growing need for the development of autonomous gait phase detection approaches to overcome these limitations. While existing automatic gait phase detection methods [9]–[12] are highly sensitive to variations in walking speed, leading to unsatisfied results across different walking speeds [13].

In this paper, we aim to develop a deep learning-based gait phase estimation model for automatic and robust recognition

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP20KK0256.

Yan Guo, Dai Owaki and Mitsuhiro Hayashibe are with the Department of Robotics, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan (e-mail: guo.yan.p6@dc.tohoku.ac.jp; owaki@tohoku.ac.jp; hayashibe@tohoku.ac.jp).

Yonatan Hutabarat is with the Hertz Chair for Artificial Intelligence and Neuroscience, University of Bonn, Bonn, Germany(e-mail: y.hutabarat@uni-bonn.de).

of gait phases across multi-scene ambulation and different walking speeds based on the use of Temporal Convolutional Network model (TCN). The main contributions of this paper are summarized as follows:

1) In order to achieve automatic and robust gait phase estimation across multi-scene ambulation and different walking speeds, we proposed a temporal convolutional network (TCN) based approach and implemented six machine learning models and two neural network models as benchmarks for comparison.
2) Extensive experiments on both real-world overground walking data and public treadmill datasets prove the superior performance in terms of estimation accuracy and robustness of the proposed method.
3) To our knowledge, few existing studies concern the comparison of robustness to speed variations in deep learning-based phase estimation methods, we design a novel evaluation method by training and validation models with different speeds for better verifying speed robustness of the deep learning model.

The rest of this paper is organized as follows. Section II introduces the related works in the gait phase detection field. Section III gives the details of the experimental protocol, along with the data acquisition and labeling processes. Section IV presents the gait phase recognition method and benchmarks. Experimental results are presented in Section V and discussed in Section VI. Section VII concludes this paper.

II. Related Works

A. Gait Analysis Systems and Sensing Modalities

1) Force: Utilizing force information to detect gait phases was the most straightforward and convenient method in the gait analysis field. With advances in electronic technology, a variety of force-based automated instruments for gait measurement have been developed, some famous golden standard in the field including AMTI force plates (FPs) [14], GAITRITE pressure sensitive walkway [15], etc. Researchers employ these devices in laboratory settings to collect the necessary force information, and it is also the most accurate method and is now serviced as the ground truth for phase detection [16]. But meanwhile, these devices are restricted to laboratory or clinical settings and are costly for daily usage. With the development of wearable technology, force-sensitive resistor (FSR) [17] and other flexible pressure insoles [18] has been used to perform gait analysis by measuring the ground reaction forces exerted on lower limbs. However, these methods have problems such as short service life. Moreover, the performance of flexible resistive materials is easily affected by temperature, resulting in inaccurate measurement outcomes.

2) EMG: Apart from the direct force information, surface Electromyography (EMG) data, which measures the continuous contraction and relaxation of the target muscles from electrode-skin interfaces, is a significant indicator of the action potential and can provide a kinetics clue of human walking patterns [19]. Recent studies have increasingly focused on employing EMG data for various motion analysis tasks, such as hand gestures [20] and gait phase recognition [21], [22]. However, raw EMG signals are inevitably affected by several factors, including noise, motion artifacts, skin-electrode interface, and cross-talk, resulting in unsatisfactory outcomes in distinguishing similar but different motions [23]. Additionally, individual differences are exacerbated by the different placement of wearable sensors and variations in muscle conditions across trials and subjects.

3) Visual Methods (RGB/RGBD): In the gait analysis field, multi-camera systems for motion capture are considered the gold standard solutions, which obtain the high-precision kinematic data by accurately detecting 3D positions of reflective markers affixed to the anatomical landmarks of subjects [24]. Some commercial Mocap systems, like Vicon [25] and Optitrack [26], are widely used in prominent biomechanics laboratories. However, due to the high cost, complicated equipment settings, and tedious preparation process, the Mocap system is usually restricted to laboratories and clinical settings. Recent advancements in computer vision have facilitated the 2D or 3D skeleton estimation directly from RGB or RGBD images. Several techniques have been proposed to achieve gait phase recognition from the estimated 3D kinematic data [27], [28]. However, inherent challenges such as 3D skeleton estimation accuracy and occlusions persist. It is still challenging to apply for gait phase estimation due to the unsatisfactory estimation of foot movements in many daily life scenarios [29].

4) Inertial Measurement Units (IMUs): Functions like multi-camera Mocap systems, IMUs also can measure the kinematic data by attaching the sensor to the corresponding body parts of interest. The methods based on the measurement of the IMU signals present several advantages over traditional Mocap systems, such as portability, low energy consumption, low cost, durability, and suitability for use outside laboratory settings, etc. Thus, many researchers begin to use inertial measurement units (IMUs) to enable gait analysis in daily life scenarios [30]–[34].

B. Heuristic methods for phase segmentation

Most existing IMU-related methods rely on heuristic approaches. These methods use hand-crafted feature extractions, such as applying thresholds, local minimum/maximum, filtering algorithms, or zero-crossing detection. Aminian et al [35] utilized the periodic variations in sagittal plane gyroscope angular velocity during the walking cycle to identify two critical gait events, heel-strike (HS) and toe-off (TO). The local minimum on either side of the angular velocity peaks served as features for recognizing these key events, thus dividing the gait phase into the swing phase and stance phase. Kotiadis et al [36] developed a finite state machine representing the gait process and implemented threshold-based criteria to determine the rules governing state transitions. Gouwanda et al [37] employed the zero-crossing technique on gyroscopic data to detect HS and TO accurately. These methods are simple and effective, as they only need to find some signal features related to certain gait events to obtain the corresponding results. These methods have greatly promoted the development of IMU-based gait analysis. However, due to the high sensitivity of accelerometer and gyroscope signal features to variations

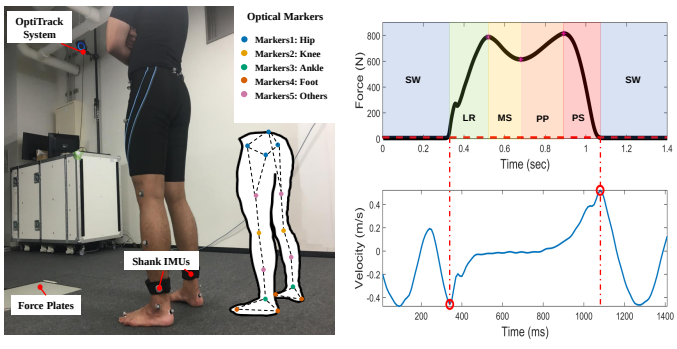
in gait patterns and walking speed, the application of these heuristic methods often requires field-specific expertise to adjust the thresholds every time in order to achieve the desired results.

C. Learning-Based gait phase estimation

1) Machine Learning: To address the issue arising from heuristic methods, researchers have proposed using learning-based gait phase estimation. Firstly, researchers use machine learning algorithms to automatically detect and classify different gait phases. Perez-Ibarra et al. [38] employed support vector machine (SVM) techniques to analyze data collected from a single-IMU foot-mounted wearable device during overground and treadmill walking sessions, involving both healthy individuals and Parkinson's disease patients, for the purpose of accurately identifying gait events. Attal et al. [39] implemented a multiple-regression hidden Markov model (HMM) to facilitate the automatic recognition of gait phases. However, this approach has some limitations as the learning ability of machine learning algorithms is limited by the complexity of the rules or features. The broader class of machine learning systems relied heavily on feature engineering, which is to transform the input data into hand-crafted features and then facilitate the downstream machine learning algorithms to solve the specific tasks. As a result, the accuracy of the results obtained using these algorithms may be limited.

2) Deep Learning: To overcome these limitations, researchers are exploring the use of deep learning algorithms. Deep learning ways can learn complex patterns and relationships in gait data without the need for pre-defined rules or features. Wang et al. [40] implemented a convolutional neural network (CNN) based algorithm to analyze data from plantar pressure sensor arrays and acceleration measurements, with the objective of accurately recognizing gait phases. Arshad et al. [41] employed a CNN-RNN hybrid model in conjunction with a single waist-mounted IMU sensor to accurately predict gait events. Su et al. [42] carried out a study with 12 healthy participants, employing 7 IMUs and 4 foot switches for the collection of walking data. They subsequently implemented a long short-term memory (LSTM) neural network to predict both gait trajectories and gait phases, demonstrating the efficacy of this approach. These deep learning-based approaches have shown promising results in gait phase detection and have significant potential for improving the accuracy and reliability of gait analysis.

In this paper, we target at learning-based gait phase estimation, which aims to achieve a model that robustness to variations in gait patterns and walking speed.



(a) (b)

Fig. 1. (a) Experimental setup. (b) Validation of data pre-processing and illustration of phase labeling.

III. EXPERIMENTS AND DATA COLLECTION

In the experimental section of our study, we utilized two distinct sets of data, encompassing both overground walking and treadmill walking scenarios, to comprehensively assess the performance of the proposed model.

A. Experimental Setup

In the overground walking experiment, 2 IMU sensors (MetaMotionC, MbientLab, San Fransisco, CA, USA) were used to obtain gait data as input to subsequent models. Owing to the high accuracy of Force Plates (FPs), FPs now serve as the gold standard for gait event detection in the gait analysis field. We use a force plate (AMTI, MA, USA) sampled at 1000 Hz to provide foot ground reaction force (GRF) as a benchmark to assess the performance of the proposed method.

Five subjects with no history of severe lower-limb related injuries participated in this study. All subjects wore two wireless IMU sensors mounted on the shank with velcro belts. Accelerometer and gyroscope data with 100 Hz sampling rate were logged in the sensors during the trials of this experiment and later uploaded to the computer for further data processing. The gyroscope range is set at 1000°/s, and the accelerometer range is set at 8g. Subjects were asked to walk back and forth on an 8m sidewalk at their preferred speed, fast walking and slow walking separately. Each experiment was conducted for 3 minutes without intervention, for a collection of total 9 minutes of gait data per subject.

In this experiment, the motion capture system (Mocap) and force plate (FP) are well synchronized through the electronic synchronization device. In order to synchronize the data of the IMUs with a force plate, participants were asked to perform the corresponding landmark actions before starting the walking session and after the end of the last walking cycle (In this experiment, participants were instructed to jump). Figure1.a depicts the scenarios of the experimental setup.

B. Biomechanical Dataset

In the session on treadmill walking, we utilized an open-source dataset, which can be found in [43]. This dataset comprised data from 15 healthy subjects (8 males and 7 females) walking on an instrumented treadmill in three different speed conditions: PWS (preferred walking speed: 3.9 ± 0.5 km/h), PWS+20%, PWS-20%. Each condition contained at least 120 steps. Reference data for this dataset were recorded using an OptoGait system (Microgate, Bolzano, Italy) and FDM-THQ pressure distribution measurement system (Zebris Medical GmbH, Isny, Germany), both integrated into the quasar® med treadmill (h/p/cosmos sports and Medical GmbH, Nussdorf-Traunstein, Germany). The Zebris system

consists of a capacitive force sensor integrated directly under the treadmill to provide ground reaction force (GRF) data from the foot, thereby offering data on the stance period. The OptoGait system consists of infrared LED transmitting and receiving bars integrated into the treadmill footboards. During data collection, the OptoGait LED sensors measure contact time and position, thereby automatically providing spatiotemporal gait parameters for treadmill walking after data collection. The dataset also includes IMU data recorded using seven Physilog®5 IMUs (Gait Up, Lausanne, Switzerland). The IMU sensors are factory-calibrated, with all sensors being well synchronized. The zebris system and IMU are configured to record at 128 Hz, while the OptoGait system is set to record at 1000 Hz. The gyroscope range is set at 1000°/s, and the accelerometer range is set at 16g. Additionally, this dataset offers video recordings from two RGB cameras capturing the frontal and sagittal views of the participants' lower bodies during the experiment. In this study, we only use two shank-mounted IMUs in this dataset.

C. Data Pre-processing

During the overground walking experiment, we collected the collected data of the participants including marker positions, shank IMUs signals, and ground reaction force information (GRF) on the force plates, firstly we use a boolean stance flag $H(t)$ to distinguish the standing period from the GRF information, the formula is shown in Eq.(1), where $F_{GRF}(t)$ is the GRF in the vertical direction from the FP, $F_{threshold}$ is the set judgment threshold, and in this study the threshold is set to 10N. The heel-strike event is then detected from the rising edge of the stance flag $H(t)$.

$$H(t) = Sgn(max(F_{GRF}(t) - F_{threshold}, 0)) \tag{1}$$

In these steps without GRF information during the overground walking experiment, the toe-off and heel-strike events were extracted by applying Foot Velocity Algorithm (FVA) [44] to the corresponding foot marker's kinematics information. We utilized this method to complement gait event data for overground walking that were not captured by force plates, which allowed us to enhance the comprehensiveness of our analysis. After applying FVA, we extend valid steps from 424 steps on the force plate to 2505 steps extracted from kinematic data. Figure1.b shows the process of verifying the effectiveness of the FVA algorithm.

Regarding the data labeling process, we initially collected all maximal and minimal values of the vertical ground reaction force (GRF) from the 424 steps on force plates. We then calculated the statistical averaging percentage of each gait phase throughout the entire support period for these steps from FP. Next, we aligned each kinematic extracted step's timestamps with 6D IMU data to create the respective data segments, then applied the phase statistical percentage obtained from the above-mentioned steps to the 2505 ground walking data segments, obtaining the corresponding phase change timestamps for each segment. Finally, we labeled the segment with each phase according to the corresponding timestamps or the closest one, thus dividing them into load response (LR), midstance (MS), propulsive phase (PP), pre-swing (PS) and swing (SW) 5 phases. Figure1.b depicts a schematic of phase labeling.

In the treadmill walking scenarios, the reference system data provided by the dataset gives the temporal information of each step of the subjects so that we can segment each step and divide it into load response (LR), midstance (MS), propulsive phase (PP), pre-swing (PS) and swing (SW) 5 phases, consistent with the phases we labeled with the overground walking data.The final format for each segment in both datasets is 8 $\times$ S (8: timestamps + 6D IMU signals + labels, where S is the number of samplings in one step = sample rate of IMU * step duration)

IV. Methods

A. Problem Statement

Before going into the implemented sequence model and the benchmarks, we first give an overview of the pipeline for learning-based phase estimation. Considering that we have N_t pre-processed training subject datasets $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \cdots, \mathcal{D}_{N_t}\}$, the main objective of the learning-based method is to establish a map function $f(\mathbf{X}_{input}, W_p)$, where W_p was referred to as the weights parameters and the set of variable $\mathbf{X}_{input} = \{\mathbf{X}_1, \cdots, \mathbf{X}_n\}$ is the input space which was described in Eq.(2).

$$\mathbf{X}_{input} = \begin{bmatrix} x_1[1] & \cdots & x_f[1] \\ \vdots & \ldots & \vdots \\ x_1[n] & \cdots & x_f[n] \end{bmatrix}, \mathbf{X} \in \mathbb{R}^{N_s \times F} \tag{2}$$

N_s stands for the sum of the sampling number of each step segment, F stands for the number of features, and the input feature in this research is 3D acceleration adding 3D angle velocity, 6 features in total. The output space $\hat{\mathbf{Y}}_p$, which denoted the output predicted phases in this study, was described in Eq.(3).

$$\hat{\mathbf{Y}}_p = \begin{bmatrix} \hat{y_1} \\ \vdots \\ \hat{y_n} \end{bmatrix}, \hat{y_n} \in [0, 1, 2, 3, 4] \tag{3}$$

$\hat{y_n}$ ranges from 0 to 4, as it corresponds to the target phases, which are divided into five stages in this study. The values 0 to 4 each represent the target values for these five phases in this study. During the training process, function spaces $\mathcal{F}$ with different learning-based algorithms $f \in \mathcal{F}$ extract the input data from provided training dataset $\mathcal{D}$ and transform it into an optimization problem: minimize loss function $\mathcal{L}$.

B. Implemented Sequence Model and Benchmarks

Two types of gait phase estimation approaches were implemented in this study: the machine learning (ML) baseline method and the deep learning (DL) baseline method.

ML Baseline Methods - The machine learning baseline methods include some classic feature extraction-based classification ML algorithms to extract time series features into gait phases, i,e., K-nearest neighbor (KNN), naive Bayes (NB),

support vector machine (SVM), decision tree (DT), logistic Regression (LG), multilayer perception (MLP), etc.

FCN - Fully convolutional network (FCN) was selected as a benchmark neural network for comparison with our methodology, given its proven success in estimating the gait phase in previous research [45], [46]. We implemented an FCN that uses a series of fully connected layers comprising hidden nodes activated by a nonlinear activation function to estimate the output. Based on our hyperparameter optimization, the FCN in our study is made up of three hidden layers activated by the ReLU function and one output layer. A dropout follows each hidden layer to aid the model in learning more complex representations and preventing overfitting.

LSTM - Long short-term memory network (LSTM), a class of recurrent neural networks, served as another comparison benchmark neural network to our method and shows superiority in most time series tasks recently, including the gait phase estimation [42]. LSTM is composed of three gates and two states: input gate i_t, forget gate f_t, output gate o_t, cell state c_t, and hidden state h_t.

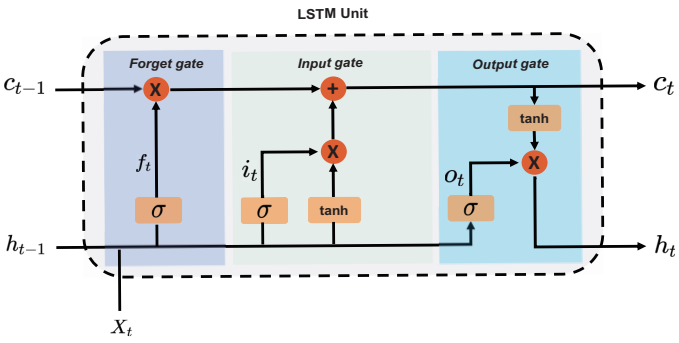


Fig. 2. Schematic Diagram of LSTM Unit.

The input gate determines which elements of the new data are retained in the cell state. The forget gate layer decides the portions of information that should be omitted from the cell state. The output gate ascertains the quantity of information on the cell state at the current time adopted as the output, cell state and hidden state serve as repositories for long-term and short-term information. The schematic representation of an LSTM unit is depicted in Figure.2, and the mathematical expressions of LSTM cell units from the input to the output can be expressed as follows:

$$i_t = \sigma\left(W_{ii}x_t + b_{ii} + W_{hi}h_{(t-1)} + b_{hi}\right) \tag{4}$$

$$f_t = \sigma\left(W_{if}x_t + b_{if} + W_{hf}h_{(t-1)} + b_{hf}\right) \tag{5}$$

$$g_t = \tanh\left(W_{ig}x_t + b_{ig} + W_{hg}h_{(t-1)} + b_{hg}\right) \tag{6}$$

$$o_t = \sigma\left(W_{io}x_t + b_{io} + W_{ho}h_{(t-1)} + b_{ho}\right) \tag{7}$$

$$c_t = f_t * c_{(t-1)} + i_t * g_t \tag{8}$$

$$h_t = o_t * \tanh\left(c_t\right) \tag{9}$$

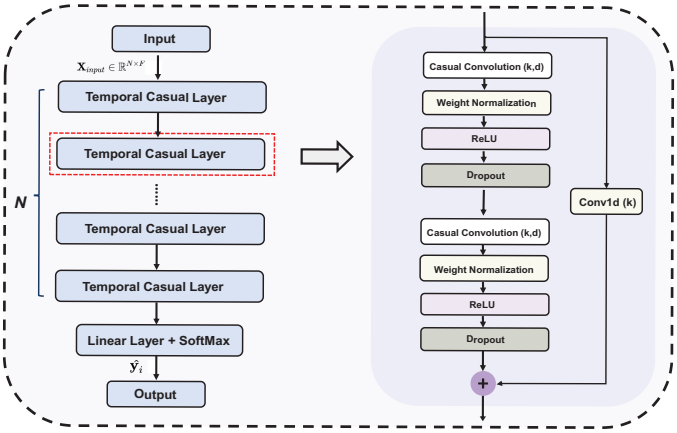


Fig. 3. Schematic Diagram of TCN Unit.

In this study, the hyperparameter-optimized LSTM we use comprises two LSTM layers each of size 128, and one fully-connected layer. It takes in six features to generate predictions for five distinct phases.

TCN (This Work) - Temporal convolutional network (TCN) introduced by Bai et al [47] was implemented as our candidate network for estimating the gait phase due to the benefits of dilated, causal convolution operation, which allowed to apply a flexible receptive field size and the convolution is done in parallel without waiting for the previous timesteps compared to a conventional convolutional network. Fig.3 shows the diagram of a TCN unit. The dilated convolution operation F on elements j for the one-dimensional input sequence $\boldsymbol{\beta}$ with the filter $\psi : \{0, ..., k-1\} \rightarrow \mathbb{R}$ is defined in Eq.(10), where d is the dilation factor, k is the filter size and $j - d_i$ indicates the direction of the past.

$$F_j(x) = \sum_{i=0}^{k-1} \psi(i)\beta_{j-d_i} \tag{10}$$

After the causal convolution operation, weight normalization is performed and then the ReLU function is used for activation, and finally perform dropout to avoid overfitting. For the purpose of enhancing the network's performance and stabilizing the model, a low-kernel convolution was added to the input tensor $\mathbf{x}$ as a residual block. The output o of the temporal causal layer is related to the input via residual block connection, which is a branch leading out to a series of transformations f:

$$o = Activation(\mathbf{x} + f(\mathbf{x})) \tag{11}$$

Finally, the receptive field size R of this network depends on the size of the filter and the dilation factor d, which was usually exponentially increased by an i times exponential relationship of 2, i represents the number of temporal causal layers in the unit.

$$R = 1 + \sum_{i=0}^{k-1} 2(k-1)d_i, d_i = 2^i \tag{12}$$

In this study, the number of temporal causal layers N was set to 3 according to our hyperparameter results. We

implemented a TCN composed of three stacked temporal causal layers with a filter kernel size of 64.

All these model was implemented in Pytorch framework. The hyperparameters of the FCN, LSTM, and TCN were independently optimized using grid search hyperparameters. Each neural network was trained using the Adam optimizer and a CrossEntropy loss function(described in Eq.(13)). The optimized models for each type of neural network were then trained from random initialization. The results for all further analyses were computed using the same training method.

$$\mathcal{L}_c = -\sum_{i=1}^{N} y_{o,i} \log p_{o,i} \tag{13}$$

C. Evaluation Metrics

For the purpose of quantified evaluation of these model's performance, in this study, the performances of these implemented gait phase estimation models are evaluated using four metrics: accuracy, precision, recall, and F-score, which can be calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{P} \tag{16}$$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively.

$$F_\alpha \cdot score = \frac{(1 + \alpha^2) \cdot Recall \cdot Precision}{\alpha^2 \cdot Recall + Precision} \tag{17}$$

In this study, α is set to 1 to assign the same weight to recall and precision, obtaining the F1-score metric.

D. Statistical Analysis

To indicate statistical significance across different networks in the robustness of different walking speeds, the mean squared error (MSE) between the predicted result and ground-truth label, which is defined in Eq.(18) was used as a compact measure for comparison metrics. A one-way analysis of variance (ANOVA) with a significance level α set to 0.05 was conducted. For a comparison between pairwise baselines, a Tukey's Honest Significant Difference (HSD) post hoc correction with the same α assigned of 0.05 was applied.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{18}$$

V. RESULTS

A. Validation of FVA and Label Accuracy

To validate the reliability of the FVA algorithm used in this study and the accuracy of the phase label method, the 424 steps collected from the force plate serves as the gold standard for validation of both FVA algorithm and phase label during stance phase.

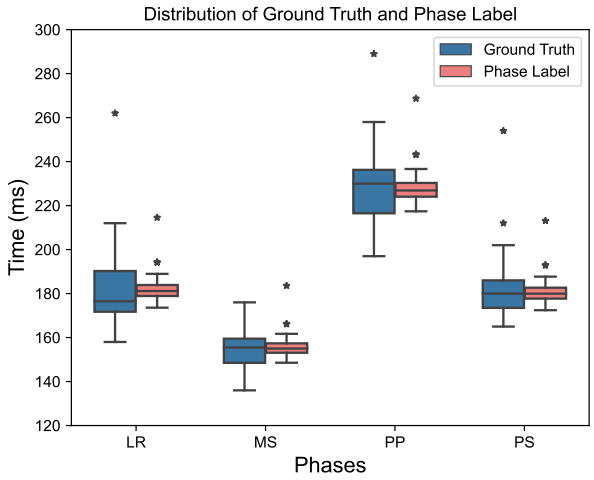


Fig. 4. Boxplot of distribution of labels derived from ground truth and the phase label in overground walking dataset

1) Validation of FVA Algorithm Derived Gait Events: Table I shows the difference between the results of the FVA algorithm and the true value of the GRF for heel-strike and toe-off detection.

TABLE I
VALIDATION OF FVA ALGORITHM

Gait Events	Validation of FVA Algorithm (Mean±Std)		
	RE	AE	ME
HS	2.014±9.588	7.957±5.640	8.000
TO	0.286±21.422	17.943±11.505	15.000

Note: HS: Heel strike, TO: Toe Off, RE: Relative Difference Error, AE: Absolute Error, Me: Median of Absolute Error. Unit in this Table: millisecond(ms).

The RE and AE of heel-strike (HS) and toe-off (TO) are within 15ms and 30ms respectively, which is consistent with the results given in the existing study [44]. The above results show that the gait events extracted from the kinematic data by FVA used in this study are consistent with the force plate reference standard, which means that the FVA algorithm can be used to accurately extract gait events of participants during walking sessions from Mocap kinematic data.

2) Validation of Label Accuracy During Stance Phase: Since the determination of HS and TO events have clearly defined the stance phase and swing phase, we only need to verify the accuracy of the remaining four labels inside the stance phase. Our label reference for the stance phase comes from the gold standard: the steps on the force plate. We use statistical values of these reference data to label all steps, the data distribution of the phase label and ground true value is shown in Figure 4 and the validation accuracy results are shown in Table II. The accuracy results of each label inside the stance phase are within 20ms.

B. Estimate Results of Phase Models

For the purpose of quantified evaluation of these models' performance, we initially partitioned the training and vali-

Fig. 5. The comparison of confusion matrix results among phase estimation models

TABLE II
VALIDATION OF STANCE PHASE LABELS

Gait Phases	Accuracy of Phase Labels (Mean±Std)		
	RE	AE	ME
LR	0.013±14.143	10.780±9.060	9.070
MS	-1.038±11.555	9.059±7.164	7.785
PP	0.035±14.958	11.708±9.198	10.344
PS	1.075±10.556	7.303±7.647	5.043

Note: RE: Relative Difference Error, AE: Absolute Error, Me: Median of Absolute Error. Unit in this Table: millisecond(ms).

dation sets at the same speed in the same subject and used the metrics such as accuracy, precision, recall, and F1-score mentioned above to evaluate.

Within the gait data of the same speed for the same subject, the training and validation sets were divided into 70% and 30%, respectively. During the training and validation, we trained and validated each speed for each subject separately. We then combined the predicted labels from these models with the true labels from all subjects for the final model evaluation.

The overall performances of phase models of ML and DL baseline methods are compared in Table III. For the Walking Session, deep learning models LSTM and TCN outperformed the other models across all evaluation metrics. Specifically, TCN achieved excellent results with accuracy, precision, recall, and F1 score of 0.9694, 0.9697, 0.9691, and 0.9694, respectively. LSTM also demonstrated high performance with accuracy, precision, recall, and F1 score of 0.9583, 0.9602, 0.9579, and 0.9583, respectively. For the Running Session, Similarly, the deep learning models LSTM and TCN also perform well, all evaluation metrics are higher than other models. TCN performed best, with accuracy, precision, recall, and F1 scores of 0.9713, 0.9716, 0.9713, and 0.9712, respectively. The

TABLE III
COMPARISON OF OVERGROUND WALKING AND TREADMILL WALKING TEST SET PERFORMANCE ACROSS ALGORITHMS

		Overground Walking Session				Treadmill Walking Session			
	Algorithms	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
ML	KNN	0.9183	0.9183	0.9179	0.9173	0.9353	0.9339	0.9319	0.9328
	NB	0.8669	0.8673	0.8670	0.8570	0.8677	0.8662	0.8542	0.8533
	SVM	0.9052	0.9051	0.8993	0.9017	0.9074	0.9092	0.8979	0.9019
	DT	0.9195	0.9208	0.9197	0.9195	0.9271	0.9240	0.9229	0.9234
	LG	0.8680	0.8680	0.8683	0.8683	0.8695	0.8657	0.8526	0.8573
	MLP	0.9433	0.9431	0.9433	0.9433	0.9465	0.9470	0.9412	0.9438
DL	FCN	0.9438	0.9445	0.9440	0.9444	0.9460	0.9463	0.9460	0.9461
	LSTM	0.9583	0.9602	0.9579	0.9583	0.9665	0.9669	0.9665	0.9664
	This Work	**0.9694**	**0.9697**	**0.9691**	**0.9694**	**0.9713**	**0.9716**	**0.9713**	**0.9712**

*Note: Numbers are the performance of each model's evaluation metrics in the training and validation **in the same subject and same speed situation**. The number in **bold** represents the highest model performance.*

evaluation metrics of LSTM are also above 0.96, which is an excellent performance. Among the machine learning models employed in these two sessions, the MLP model exhibited the best performance, with all evaluation metrics above 0.94. The SVM, KNN, and DT models also demonstrated comparatively good performance. In general, however, in the gait phase estimation tasks among these two scenarios, the deep learning models outperformed the machine learning models.

Figure.5 illustrates the confusion matrix resulting from the model evaluation that combined the predicted labels from models trained and validated on data from the same subject at the same speed, with the true labels from all participants. The confusion matrix provides details of the estimation accuracy of each model at each phase. Consistent with the results shown in Table I, the deep learning models LSTM and TCN outperformed other models in estimating most phases with higher precision. To offer a more intuitive display of the differences in the results of these models, Figure.6 presents an example, displaying a comparison between the predicted label and true labels for phase estimation by these models at the same speed for the same subject.

In conclusion, the deep learning models LSTM and TCN excelled in phase estimation tasks during both the overground walking session and treadmill walking session, while the TCN slightly outperforms the LSTM.

C. Inter-Subject Model Generalizability Evaluation

Based on the conclusions from previous estimation performance evaluation and considering that deep learning models generally handle dynamical time-series data better than machine learning models in most cases [48], we are confining our subsequent evaluation to the deep learning models: FNN, LSTM, and TCN.

To assess the generalizability of these models, we employ the Leave-One-Subject-Out (LOSO) cross-validation method. In this study, we utilized the data of different subjects at the same speed to train the model, while data from one subject is left out for validation. We averaged the results obtained after validating the data of all subjects in turn to obtain the final results. Table IV and Table V summarize the final results in normal, fast, and slow speed situations of each model after LOSO in the overground walking session and treadmill walking session respectively.

As the table shows, whether in the overground walking session or the treadmill walking session, the accuracy at slow speed is consistently higher than that in both fast and normal scenarios, while the lowest accuracy is observed at fast speed. This can be explained by gait features that might be more easily captured by neural networks for slow motion. For the LOSO results of each network, TCN has achieved the best results compared to FCN and LSTM in both the overground walking session and the treadmill walking session. The overall trend for FCN and LSTM is similar for this inter-subject validation, this indicates that LSTM has some weakness for cross-subject usage.

Overall, in terms of LOSO results, TCN consistently outperforms other benchmarks. These results demonstrate the robustness of the TCN model in handling cross-subject estimations. The certain advantage over FNN and LSTM indicates the TCN model's generalizability for different subjects, which is an important feature when we need gait phase estimation for a brand-new subject.

D. Intra-Subject Speed Variation Robustness

To evaluate the speed variation robustness of these models, we conducted training and validation across different speeds for the same subject.

1) Evaluation of Speed Variation Robustness: In this study, we trained the model using data from one speed, while

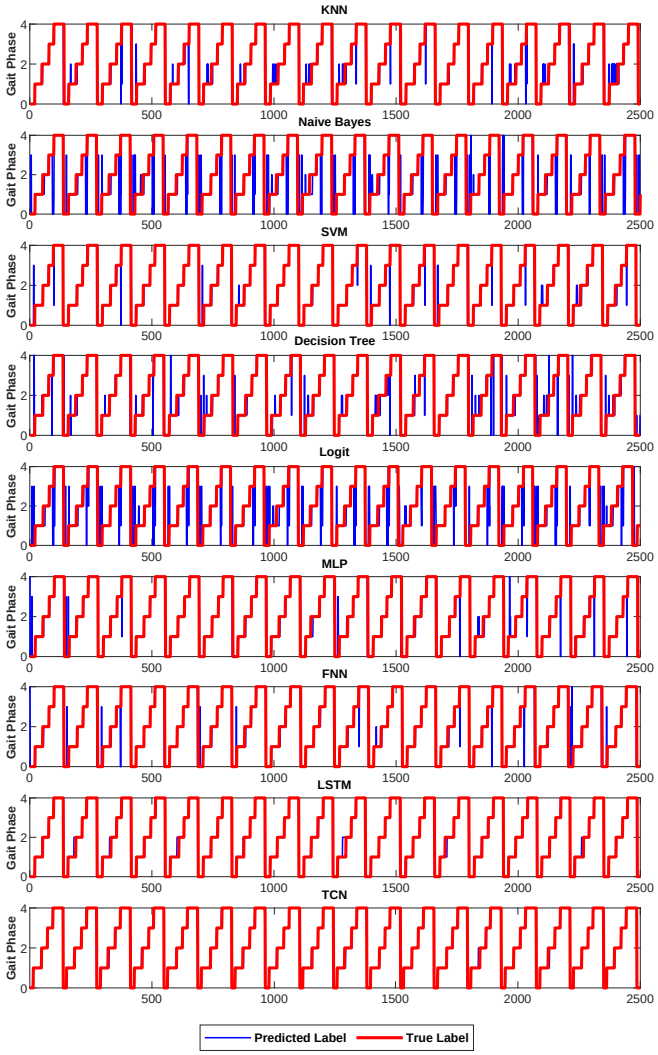


Fig. 6. Comparison figure of predicted labels versus true labels across all algorithms. The blue line denotes the predicted label, and the red line denotes the truth label (0-4 in the y-axis denotes the label of target phases).

validating it on data from another speed. After sequentially validating all subjects' data, we selected all valid results for analysis (Due to the absence of effective slow-speed data from one participant in the treadmill dataset, we excluded this individual from training and validation processes involving slow-speed data). Then we averaged the obtained results to yield a final outcome. Table VI and Table VII summarize the final results for each model.

Table VI shows the results of the validation accuracy of each neural network in the speed variation robustness evaluation. Whether in the overground walking session or the treadmill walking session, TCN consistently demonstrates the highest accuracy, with the accuracy exceeding 90% in most cases. While the LSTM exhibits some advantages over the FCN but is worse than TCN. Compared with the three neural network results of training and validation in the same subject and same speed situation in Table III, the accuracy of the FCN network drops significantly. Although the accuracy of the

TABLE IV
INTER-SUBJECT MODEL GENERALIZABILITY RESULT OF EACH NEURAL NETWORK IN OVERGROUND WALKING SESSION

	Overground Walking Session		
Networks	Normal	Fast	Slow
FCN	0.6937±0.077	0.6626±0.081	0.7258±0.071
LSTM	0.6923±0.074	0.6653±0.072	0.7342±0.079
This Work	**0.7693±0.082**	**0.7364±0.077**	**0.8032±0.071**

*Note: Numbers are the average test accuracy of leave-one-subject-out (LOSO) results of each participant in an overground walking session. The number in **bold** represents the highest generalizability performance.*

TABLE V
INTER-SUBJECT MODEL GENERALIZABILITY RESULT OF EACH NEURAL NETWORK IN TREADMILL WALKING SESSION

	Treadmill Walking Session		
Networks	Normal	Fast	Slow
FCN	0.6820±0.074	0.6556±0.079	0.7090±0.065
LSTM	0.6687±0.079	0.6341±0.072	0.7105±0.076
This Work	**0.7607±0.078**	**0.7101±0.089**	**0.7814±0.069**

*Note: Numbers are the average test accuracy of leave-one-subject-out (LOSO) results of each participant in a treadmill walking session. The number in **bold** represents the highest generalizability performance.*

LSTM in Table III is relatively close to TCN, there is an evident difference between the results between LSTM and TCN in Table VI.

While accuracy can only measure the absolute correct rate of the model's estimation, to quantify the distance difference between the predicted labels and the true labels of each neural network model in the speed variation robustness test, we calculated the mean square error (MSE) of each model in various training and validation situations. As shown in Table VII, for the overground session, across all different speed training-validation configurations, the TCN's MSE is the lowest. In comparison to FCN, the average MSE of TCN in various speed training-validation scenarios has been reduced by 55.93%, 64.40%, 49.75%, 48.87%, 65.59%, and 54.76%, respectively. Against LSTM, it has been reduced by 46.99%, 51.22%, 38.27%, 26.15%, 60.33%, and 45.94% respectively. While LSTM also shows a decrease in MSE compared to FCN, the reduction is consistently below 30%. The greatest reductions in MSE correspond to the training-validation scenarios TSVF and TFVS, which involve the most significant differences in speed.

For the treadmill session, the MSE of TCN is again the lowest across all different speed training-validation configurations. Compared to FCN, the average MSE of TCN in various speed training-validation scenarios has been reduced by 50.19%, 59.47%, 47.61%, 55.97%, 53.22%, and 56.93% respectively. Against LSTM, it has been reduced by 44.14%, 51.73%, 42.41%, 50.32%, 50.45%, and 46.26% respectively. Similarly, LSTM shows a decrease in MSE compared to FCN, but the performance is mostly similar, with the MSE difference

TABLE VI
COMPARISON OF SPEED VARIATION ROBUSTNESS WITH ACCURACY EVALUATION OF EACH NEURAL NETWORK

	Overground Walking Session						Treadmill Walking Session					
Networks	TSVN	TSVF	TNVS	TNVF	TFVS	TFVN	TSVN	TSVF	TNVS	TNVF	TFVS	TFVN
FCN	0.811(0.036)	0.758(0.039)	0.812(0.030)	0.819(0.032)	0.798(0.028)	0.859(0.031)	0.838(0.038)	0.762(0.031)	0.824(0.029)	0.841(0.031)	0.803(0.037)	0.839(0.03)
LSTM	0.864(0.028)	0.809(0.028)	0.866(0.027)	0.878(0.023)	0.802(0.030)	0.871(0.029)	0.867(0.026)	0.810(0.035)	0.859(0.026)	0.878(0.021)	0.798(0.032)	0.873(0.027)
This Work	0.911(0.017)	0.902(0.022)	0.902(0.018)	0.904(0.019)	0.911(0.017)	0.914(0.016)	0.911(0.023)	0.877(0.026)	0.904(0.021)	0.916(0.023)	0.864(0.027)	0.915(0.020)

Note: Numbers are the mean accuracy results (the higher the better) of each train and valid situation (T: training, V: validation, S: slow, N: normal, F: fast). These accuracy results were presented by mean values and their standard deviation.

TABLE VII
COMPARISON OF SPEED VARIATION ROBUSTNESS WITH MEAN SQUARE ERROR EVALUATION OF EACH NEURAL NETWORK

	Overground Walking Session						Treadmill Walking Session					
Networks	TSVN	TSVF	TNVS	TNVF	TFVS	TFVN	TSVN	TSVF	TNVS	TNVF	TFVS	TFVN
FCN	0.599(0.172)	0.837(0.180)	0.597(0.113)	0.577(0.149)	0.776(0.104)	0.515(0.116)	0.536(0.178)	0.824(0.168)	0.565(0.207)	0.527(0.158)	0.714(0.176)	0.534(0.144)
LSTM	0.498(0.149)	0.611(0.153)	0.486(0.148)	0.462(0.140)	0.673(0.171)	0.431(0.131)	0.478(0.143)	0.692(0.194)	0.514(0.156)	0.467(0.119)	0.674(0.189)	0.428(0.093)
This Work	0.264(0.074)	0.298(0.094)	0.300(0.074)	0.295(0.080)	0.267(0.068)	0.233(0.051)	0.267(0.104)	0.334(0.137)	0.296(0.109)	0.232(0.100)	0.334(0.120)	0.230(0.087)

Note: Numbers are the mean squared error (MSE) results (the lower the better) of each train and valid situation (T: training, V: validation, S: slow, N: normal, F: fast). These MSE comparisons were presented by mean values and their standard deviation. This error is regarding the gait phase label value (ranging from 0 to 4).

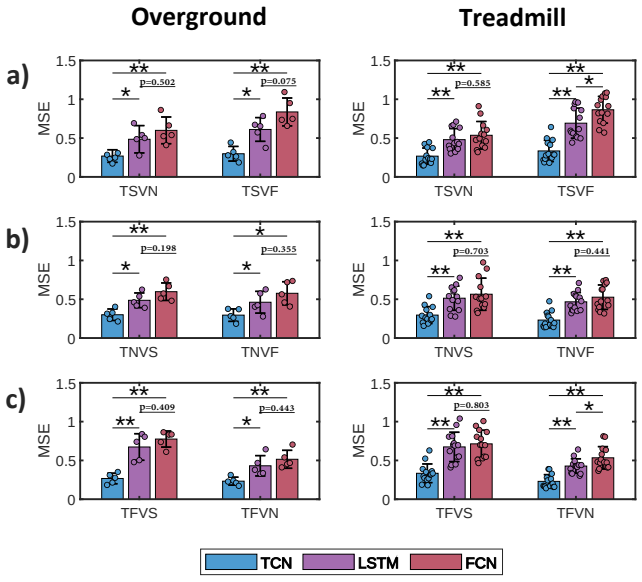


Fig. 7. Speed variation robustness results per neural network. The MSE (lower is better) results of the temporal convolutional network (TCN) and baseline methods are shown. Results are presented for the overground and treadmill ambulation modes among 12 training and validations situations (Figures a, b and c present training with slow, normal, and fast situations respectively. T: training, V: validation, S: slow speed, N: normal speed, F: fast speed). The bars represent the mean error ± the standard deviation. * represents p values ($p < 0.05$), ** represents p values ($p < 0.01$).

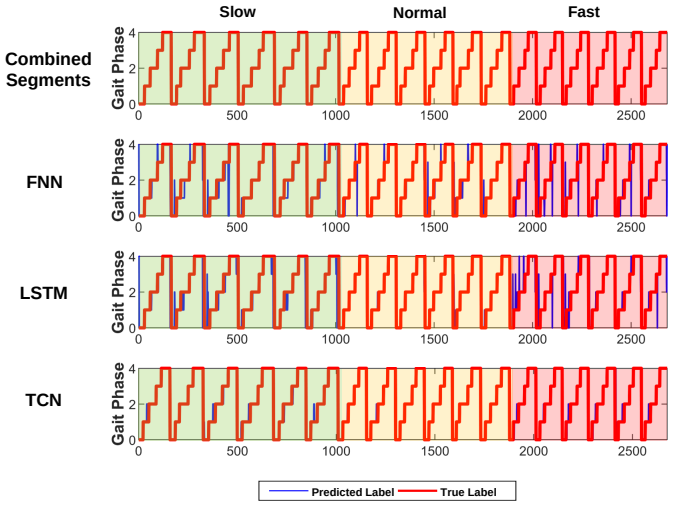


Fig. 8. Robustness performance test to train each neural network on the normal speed walking data and validate on a combined segment consisting of slow, normal, and fast speed from speed-controlled treadmill data. Each speed contained six steps, together forming a combined segment consisting of 18 steps with three speeds (0-4 in the y-axis denotes the label of target phases).

consistently below 20%. The same as mentioned above, the most significant reductions in MSE correspond to the training-validation scenarios TSVF and TFVS, which involve the most substantial differences in speed.

Fig.7 provides a more detailed illustration of the MSE differences in phase estimation among these models under various speed training-validation scenarios in the context of intra-subject conditions. The results were statistically analyzed and grouped according to different ambulation modes and training-validation scenarios. As illustrated, for the overground session, all the MSE results of the TCN model showed significant differences compared to FCN, in which the TNVF scenario exhibits $p<0.05$, while all other scenarios show $p<0.01$. All TCN results also showed significant differences compared to LSTM ($p<0.05$), where the TFVS scenario was $p<0.01$.In the case of the overground session, the results of LSTM did not

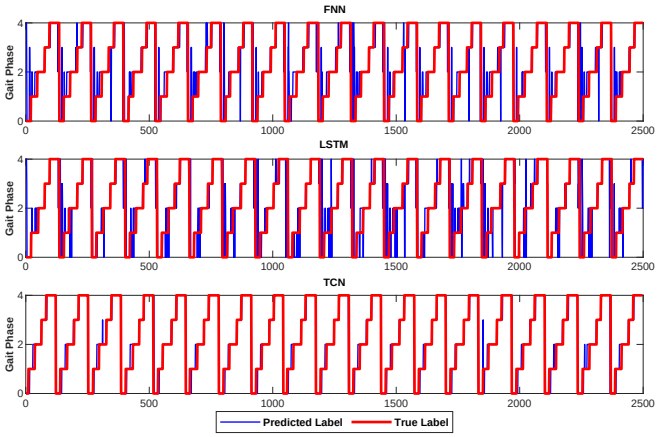


Fig. 9. One example of predicted labels versus true labels while trained with slow and validated with fast situations (TSVF). The blue line denotes the predicted label, while the red line denotes the truth label (0-4 in the y-axis denotes the label of target phases).

show a significant difference compared to FCN.

For the treadmill session, all the MSE results of the TCN model demonstrated a significant difference compared to FCN ($p<0.01$). All the results of TCN also showed significant differences compared to LSTM ($p<0.01$). In the treadmill session, the results of LSTM compared to FCN only showed significant differences in the TSVF and TFVN scenarios ($p<0.05$), and no significant differences were seen in other scenarios. All p-values for non-significant differences are marked in the figure.

2) Robustness Performance on Combined Speed: In order to understand the speed variable robustness of each network more intuitively, a robustness performance test on a combined speed segment was conducted. The combined segment consisted of slow, normal, and fast-speed walking data from the speed-controlled treadmill experiment. Each speed contained six steps, together forming a combined segment consisting of 18 steps with three speeds. Each neural network was trained in normal walking speed data (the same speed as the normal part of the combined segment) and validated on the combined segment. As we can see in Fig.8, the performance of LSTM and TCN was close in normal speed part, but TCN shows supreme robustness in the other slow and fast speed segments, which is consistent with the results shown previously.

3) Performance of Speed Robustness in Limit State: To offer an intuitive representation of the mean squared error differences in phase estimation across these models under a variety of speed training-validation scenarios, Fig.9 provides an example under the specific scenario of TSVF (training with slow and validation with fast) in treadmill walking, which is characterized by a substantial difference in speed and illustrates the comparison between the predicted labels and the true labels. It can be seen that there is a clear distinction in the results of speed variations between TCN and the other models, FNN and LSTM. This illustrates the robustness to speed variations that TCN exhibits in comparison to the other two benchmark models.

VI. DISCUSSION

According to the above results, the accuracy and robustness to different walking speeds of gait phase estimation in multi scenarios ambulation context can be enhanced by utilizing Temporal Convolutional Networks (TCN) based deep learning methods. Most existing IMU-related methods rely on heuristic approaches. These methods employ hand-crafted feature extractions, such as applying thresholds [35], identifying local minimum/maximum values [36], implementing filtering algorithms, or detecting zero crossings [37]. Due to the high sensitivity of accelerometer and gyroscope signal features to variations in gait patterns and walking speed, the accuracy and reliability of these heuristic methods are often insufficient.

In this paper, we presented a temporal convolutional network (TCN) based approach for automatic and robust recognition of gait phases across multi-scene ambulation and different walking speeds. Extensive experiments on both public datasets and real-world data have proven the superior performance of the proposed method in terms of accuracy and efficiency. We utilized the proposed TCN-based gait phase estimation model to segment the training-labeled gait series into five gait phases and employed six classical classification machine learning methods and two commonly used neural networks as benchmarks for comparison.

In the first evaluation, we split the data at the same subject of the same speed into training and validation sets and evaluated and compared with four metrics: accuracy, precision, recall, and F1 score, to select the best classifier for gait phase estimation. The results demonstrated that neural network methods outperformed other feature-based machine learning classifiers in this task, and our proposed method also surpassed the other benchmarks compared. In the second evaluation, we tested data from different subjects at the same speed, using LOSO (Leave-One-Subject-Out) to compare model generalizability. The results indicated that the TCN-based method has certain advantages over FCN and LSTM in terms of model generalization.

In the final evaluation, we assessed and compared the model's speed robustness using the data from the same subject at different speeds (To our knowledge, few existing studies concern the comparison of robustness to speed variations in deep learning-based phase estimation methods). The experimental results showed that our method exhibited significant differences in various walking conditions compared to FCN and LSTM ($p<0.05$). The differences were even more pronounced in cases of combined speed validation and significant speed variations, such as TSVF (Train with slow, validate with fast) and TFVS (Train with fast, validate with slow). In the combined speed validation, the perfomance of TCN was close to the performance of LSTM in normal speed part, this was consistent with our results in Table III. But when it comes to validation under other speed conditions, TCN showed supreme speed robustness againstother benchmarks, which is consistent with our results in the analysis of Table VII and Figure 7. The performance test in the limit state further illustrates the gap in speed robustness among TCN and benchmarks.

There are also some limitations within this study. First,

the current model is primarily designed for healthy individuals, and it can be unlikely to have the same effect on pathological gait. Therefore, consideration should be given to increasing pathological gait training data and improving the corresponding models to enhance the model's generalization ability for other gait types, enhancing it closer to real clinical gait analysis application. Second, the current algorithms and hardware systems have been tested only by offline analysis, the development of hardware and algorithms for corresponding online analysis systems can be considered in the future. Finally, compared with the RNN network structure, the parallelism processing method of TCN may struggle to capture the intricate sequential dependencies that are crucial to accurately identify the gait phase transitions, moreover where there are distinct gait patterns occur at the beginning and end of a gait cycle. We suggest looking into the multi-stage TCN [49], which uses combined large and small receptive fields and fewer parameters potentially improving the recognition accuracy in the transition phase, and other possible deep learning architectures that may offer solutions to these challenges.

VII. CONCLUSION

We presented a temporal convolutional network-based approach for automatic and robust recognition of gait phases across multi-scene ambulation and different walking speeds. We collected data on both real-world overground walking experiments and public treadmill datasets to validate the performance in terms of the accuracy and robustness of the proposed method. By comparing our method with six machine learning models and two deep neural network models, our method achieves nearly 97% accuracy in gait phase estimation for both overground and treadmill walking, outperforming all compared benchmarks. It also outperformed in both model generalizability evaluation and speed variation robustness tests over the other two neural networks. Especially, the proposed method could keep nearly 90% accuracy in gait phase estimation even when there was a significant discrepancy in speed between training data and validation data. These results demonstrate that our method has outstanding estimation performance and high robustness on gait speed variations.

REFERENCES

[1] A. Kharb, V. Saini, Y. Jain, and S. Dhiman, "A review of gait cycle and its parameters," *IJCEM International Journal of Computational Engineering & Management*, vol. 13, pp. 78–83, 2011.

[2] M. Jacquelin Perry, "Gait analysis: normal and pathological function," *New Jersey: SLACK*, 2010.

[3] I. P. Pappas, M. R. Popovic, T. Keller, V. Dietz, and M. Morari, "A reliable gait phase detection system," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 9, no. 2, pp. 113–125, 2001.

[4] M. Meghji, A. Balloch, D. Habibi, I. Ahmad, N. Hart, R. Newton, J. Weber, and A. Waqar, "An algorithm for the automatic detection and quantification of athletes' change of direction incidents using imu sensor data," *IEEE Sensors Journal*, vol. 19, no. 12, pp. 4518–4527, 2019.

[5] D. R. Howell, T. A. Buckley, R. C. Lynall, and W. P. Meehan III, "Worsening dual-task gait costs after concussion and their association with subsequent sport-related injury," *Journal of neurotrauma*, vol. 35, no. 14, pp. 1630–1636, 2018.

[6] A. Rampp, J. Barth, S. Schülein, K.-G. Gaßmann, J. Klucken, and B. M. Eskofier, "Inertial sensor-based stride parameter calculation from gait sequences in geriatric patients," *IEEE transactions on biomedical engineering*, vol. 62, no. 4, pp. 1089–1097, 2014.

[7] S. Mohammed, A. Same, L. Oukhellou, K. Kong, W. Huo, and Y. Amirat, "Recognition of gait cycle phases using wearable sensors," *Robotics and Autonomous Systems*, vol. 75, pp. 50–59, 2016.

[8] A. Muro-De-La-Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications," *Sensors*, vol. 14, no. 2, pp. 3362–3394, 2014.

[9] S. Ghoussayni, C. Stevens, S. Durham, and D. Ewins, "Assessment and validation of a simple automated method for the detection of gait events and intervals," *Gait & Posture*, vol. 20, no. 3, pp. 266–272, 2004.

[10] J. Zeni Jr, J. Richards, and J. Higginson, "Two simple methods for determining gait events during treadmill and overground walking using kinematic data," *Gait & posture*, vol. 27, no. 4, pp. 710–714, 2008.

[11] M. S. Aung, S. B. Thies, L. P. Kenney, D. Howard, R. W. Selles, A. H. Findlow, and J. Y. Goulermas, "Automated detection of instantaneous gait events using time frequency analysis and manifold embedding," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 6, pp. 908–916, 2013.

[12] Y. Hutabarat, D. Owaki, and M. Hayashibe, "Quantitative gait assessment with feature-rich diversity using two imu sensors," *IEEE Transactions on Medical Robotics and Bionics*, vol. 2, no. 4, pp. 639–648, 2020.

[13] W. Hong, J. Lee, and P. Hur, "Piecewise linear labeling method for speed-adaptability enhancement in human gait phase estimation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 628–635, 2022.

[14] AMTI, "Amti - advanced mechanical technology, inc." 2023, [Online]. Available: https://www.amti.biz/.

[15] GAITRite, "The gaitrite gold standard," 2023, [Online]. Available: https://www.gaitrite.com/.

[16] I. Papavasileiou, W. Zhang, and S. Han, "Real-time data-driven gait phase detection using ground contact force measurements: Algorithms, platform design and performance," *Smart Health*, vol. 1, pp. 34–49, 2017.

[17] C. M. Senanayake and S. A. Senanayake, "Computational intelligent gait-phase detection system to identify pathological gait," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 5, pp. 1173–1179, 2010.

[18] T. Dong, Y. Guo, Y. Gu, L. Wang, T. Liu, X. Wang, P. Li, and X. Li, "Design of a wireless and fully flexible insole using a highly sensitive pressure sensor for gait event detection," *Measurement Science and Technology*, vol. 32, no. 10, p. 105109, 2021.

[19] V. Agostini, M. Ghislieri, S. Rosati, G. Balestra, and M. Knaflitz, "Surface electromyography applied to gait analysis: How to improve its impact in clinics?" *Frontiers in neurology*, p. 994, 2020.

[20] A. Phinyomark, A. Nuidod, P. Phukpattaranont, and C. Limsakul, "Feature extraction and reduction of wavelet transform coefficients for emg pattern classification," *Elektronika ir Elektrotechnika*, vol. 122, no. 6, pp. 27–32, 2012.

[21] S. W. Lee, T. Yi, J.-W. Jung, and Z. Bien, "Design of a gait phase recognition system that can cope with emg electrode location variation," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 3, pp. 1429–1439, 2015.

[22] R. Luo, S. Sun, X. Zhang, Z. Tang, and W. Wang, "A low-cost end-to-end semg-based gait sub-phase recognition system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 1, pp. 267–276, 2019.

[23] C. J. De Luca, L. D. Gilmore, M. Kuznetsov, and S. H. Roy, "Filtering the surface emg signal: Movement artifact and baseline noise contamination," *Journal of biomechanics*, vol. 43, no. 8, pp. 1573–1579, 2010.

[24] S. Chen, J. Lach, B. Lo, and G.-Z. Yang, "Toward pervasive gait analysis with wearable sensors: A systematic review," *IEEE journal of biomedical and health informatics*, vol. 20, no. 6, pp. 1521–1537, 2016.

[25] Vicon, "Vicon motion capture," 2023, [Online]. Available: https://www.vicon.com/.

[26] OptiTrack, "Optitrack - motion capture systems," 2023, [Online]. Available: https://optitrack.com/.

[27] S. Bei, Z. Zhen, Z. Xing, L. Taocheng, and L. Qin, "Movement disorder detection via adaptively fused gait analysis based on kinect sensors," *IEEE Sensors Journal*, vol. 18, no. 17, pp. 7305–7314, 2018.

[28] Y. Guo, F. Deligianni, X. Gu, and G.-Z. Yang, "3-d canonical pose estimation and abnormal gait recognition with a single rgb-d camera," *IEEE Robotics and Automation letters*, vol. 4, no. 4, pp. 3617–3624, 2019.

[29] S. L. Colyer, M. Evans, D. P. Cosker, and A. I. Salo, "A review of the evolution of vision-based motion analysis and the integration of

advanced computer vision methods towards developing a markerless system," *Sports medicine-open*, vol. 4, no. 1, pp. 1–15, 2018.

[30] Y. Hutabarat, D. Owaki, and M. Hayashibe, "Recent advances in quantitative gait analysis using wearable sensors: a review," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26 470–26 487, 2021.

[31] L. Wang, X. Jin, Y. Sun, L. Li, Q. Li, Y. Guo, G. Cheng, and T. Liu, "Inertial sensor-based gait analysis for evaluating the effects of acupuncture treatment in parkinson's disease," in *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2019, pp. 323–328.

[32] Y. Hutabarat, D. Owaki, and M. Hayashibe, "Seamless temporal gait evaluation during walking and running using two imu sensors," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 6835–6840.

[33] Y. Wu, H.-B. Zhu, Q.-X. Du, and S.-M. Tang, "A survey of the research status of pedestrian dead reckoning systems based on inertial sensors," *International Journal of Automation and Computing*, vol. 16, pp. 65–83, 2019.

[34] Y. Wu, H. Zhu, Q. Du, and S. Tang, "A pedestrian dead-reckoning system for walking and marking time mixed movement using an shss scheme and a foot-mounted imu," *IEEE Sensors Journal*, vol. 19, no. 5, pp. 1661–1671, 2018.

[35] K. Aminian, B. Najafi, C. Büla, P.-F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *Journal of biomechanics*, vol. 35, no. 5, pp. 689–699, 2002.

[36] D. Kotiadis, H. J. Hermens, and P. H. Veltink, "Inertial gait phase detection for control of a drop foot stimulator: Inertial sensing for gait phase detection," *Medical engineering & physics*, vol. 32, no. 4, pp. 287–297, 2010.

[37] D. Gouwanda, A. A. Gopalai, and B. H. Khoo, "A low cost alternative to monitor human gait temporal parameters–wearable wireless gyroscope," *IEEE Sensors Journal*, vol. 16, no. 24, pp. 9029–9035, 2016.

[38] J. C. Perez-Ibarra, A. A. Siqueira, and H. I. Krebs, "Identification of gait events in healthy and parkinson's disease subjects using inertial sensors: A supervised learning approach," *IEEE Sensors Journal*, vol. 20, no. 24, pp. 14 984–14 993, 2020.

[39] F. Attal, Y. Amirat, A. Chibani, and S. Mohammed, "Automatic recognition of gait phases using a multiple-regression hidden markov model," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 4, pp. 1597–1607, 2018.

[40] F. Wang, L. Yan, and J. Xiao, "Recognition of the gait phase based on new deep learning algorithm using multisensor information fusion," *Sensors Mater.*, vol. 31, no. 10, pp. 3041–3054, 2019.

[41] M. Z. Arshad, A. Jamsrandorj, J. Kim, and K.-R. Mun, "Gait events prediction using hybrid cnn-rnn-based deep learning models through a single waist-worn wearable sensor," *Sensors*, vol. 22, no. 21, p. 8226, 2022.

[42] B. Su and E. M. Gutierrez-Farewik, "Gait trajectory and gait phase prediction based on an lstm network," *Sensors*, vol. 20, no. 24, p. 7127, 2020.

[43] J. Trautmann, L. Zhou, C. M. Brahms, C. Tunca, C. Ersoy, U. Granacher, and B. Arnrich, "Tripod—a treadmill walking dataset with imu, pressure-distribution and photoelectric data for gait analysis," *Data*, vol. 6, no. 9, p. 95, 2021.

[44] C. M. O'Connor, S. K. Thorpe, M. J. O'Malley, and C. L. Vaughan, "Automatic detection of gait events using kinematic data," *Gait & posture*, vol. 25, no. 3, pp. 469–474, 2007.

[45] H. T. T. Vu, F. Gomez, P. Cherelle, D. Lefeber, A. Nowé, and B. Vanderborght, "Ed-fnn: A new deep learning algorithm to detect percentage of the gait cycle for powered prostheses," *Sensors*, vol. 18, no. 7, p. 2389, 2018.

[46] Z. Wang, C. Chen, F. Yang, Y. Liu, G. Li, and X. Wu, "Real-time gait phase estimation based on neural network and assistance strategy based on simulated muscle dynamics for an ankle exosuit," *IEEE Transactions on Medical Robotics and Bionics*, vol. 5, no. 1, pp. 100–109, 2023.

[47] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[48] S. S. Tabrizi, S. Pashazadeh, and V. Javani, "Comparative study of table tennis forehand strokes classification using deep learning and svm," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13 552–13 561, 2020.

[49] S.-J. Li, Y. AbuFarha, Y. Liu, M.-M. Cheng, and J. Gall, "Ms-tcn++: Multi-stage temporal convolutional network for action segmentation," *IEEE transactions on pattern analysis and machine intelligence*, 2020.