POLITECNICO DI TORINO
Repository ISTITUZIONALE

Ecological validity of a deep learning algorithm to detect gait events from real-life walking bouts in mobility-limiting diseases

Original
Ecological validity of a deep learning algorithm to detect gait events from real-life walking bouts in mobility-limiting diseases / Romijnders, R., Salis, F., Hansen, C., Küderle, A., Paraschiv-Ionescu, A., Cereatti, A., Alcock, L., Aminian, K., Becker, C., Bertuletti, S., Bonci, T., Brown, P., Buckley, E., Cantu, A., Carsin, A., Caruso, M., Caulfield, B., Chiari, L., D(extquotesingle)Ascanio, I., Del Din, S., et al.. - In: FRONTIERS IN NEUROLOGY. - ISSN 1664-2295. - ELETTRONICO. - 14:(2023). [10.3389/fneur.2023.1247532]

Availability:
This version is available at: 11583/2983103 since: 2023-10-18T12:57:53Z

Publisher:
Frontiers

Published
DOI:10.3389/fneur.2023.1247532

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

10 November 2026

frontiers | Frontiers in Neurology

TYPE Original Research
PUBLISHED 16 October 2023
DOI 10.3389/fneur.2023.1247532

Check for updates

OPEN ACCESS

EDITED BY
Carolina Soares,
Centro Hospitalar Universitário de São João (CHUSJ), Portugal

REVIEWED BY
Diogo Carneiro,
Centro Hospitalar e Universitário de Coimbra, Portugal
Francesco Asci,
Mediterranean Neurological Institute Neuromed (IRCCS), Italy

*CORRESPONDENCE
Robbin Romijnders
✉ robr@tf.uni-kiel.de

RECEIVED 26 June 2023
ACCEPTED 18 September 2023
PUBLISHED 16 October 2023

CITATION
Romijnders R, Salis F, Hansen C, Küderle A, Paraschiv-Ionescu A, Cereatti A, Alcock L, Aminian K, Becker C, Bertuletti S, Bonci T, Brown P, Buckley E, Cantu A, Carsin A-E, Caruso M, Caulfield B, Chiari L, D'Ascanio I, Del Din S, Eskofier B, Fernstad SJ, Fröhlich MS, Garcia Aymerich J, Gazit E, Hausdorff JM, Hiden H, Hume E, Keogh A, Kirk C, Kluge F, Koch S, Mazzà C, Megaritis D, Micó-Amigo E, Müller A, Palmerini L, Rochester L, Schwickert L, Scott K, Sharrack B, Singleton D, Soltani A, Ullrich M, Vereijken B, Vogiatzis I, Yarnall A, Schmidt G and Maetzler W (2023) Ecological validity of a deep learning algorithm to detect gait events from real-life walking bouts in mobility-limiting diseases.
Front. Neurol. 14:1247532.
doi: 10.3389/fneur.2023.1247532

COPYRIGHT
© 2023 Romijnders, Salis, Hansen, Küderle, Paraschiv-Ionescu, Cereatti, Alcock, Aminian, Becker, Bertuletti, Bonci, Brown, Buckley, Cantu, Carsin, Caruso, Caulfield, Chiari, D'Ascanio, Del Din, Eskofier, Fernstad, Fröhlich, Garcia Aymerich, Gazit, Hausdorff, Hiden, Hume, Keogh, Kirk, Kluge, Koch, Mazzà, Megaritis, Micó-Amigo, Müller, Palmerini, Rochester, Schwickert, Scott, Sharrack, Singleton, Soltani, Ullrich, Vereijken, Vogiatzis, Yarnall, Schmidt and Maetzler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Ecological validity of a deep learning algorithm to detect gait events from real-life walking bouts in mobility-limiting diseases

Robbin Romijnders[1,2]*, Francesca Salis[3], Clint Hansen[2], Arne Küderle[4], Anisoara Paraschiv-Ionescu[5], Andrea Cereatti[6], Lisa Alcock[7], Kamiar Aminian[5], Clemens Becker[8], Stefano Bertuletti[3], Tecla Bonci[9,10], Philip Brown[11], Ellen Buckley[9,10], Alma Cantu[12], Anne-Elie Carsin[13,14,15], Marco Caruso[6], Brian Caulfield[16,17], Lorenzo Chiari[18,19], Ilaria D'Ascanio[18], Silvia Del Din[7,20], Björn Eskofier[4], Sara Johansson Fernstad[12], Marceli Stanislaw Fröhlich[21], Judith Garcia Aymerich[13,14,15], Eran Gazit[22], Jeffrey M. Hausdorff[22,23], Hugo Hiden[12], Emily Hume[24], Alison Keogh[16,17], Cameron Kirk[7], Felix Kluge[4,25], Sarah Koch[13,14,15], Claudia Mazzà[9,10], Dimitrios Megaritis[24], Encarna Micó-Amigo[7], Arne Müller[25], Luca Palmerini[18,19], Lynn Rochester[7,11], Lars Schwickert[8], Kirsty Scott[9,10], Basil Sharrack[26], David Singleton[16,17], Abolfazl Soltani[5,27], Martin Ullrich[4], Beatrix Vereijken[28], Ioannis Vogiatzis[24], Alison Yarnall[7,10], Gerhard Schmidt[1] and Walter Maetzler[2] on behalf of Mobilise-D Consortium

[1]Digital Signal Processing and System Theory, Electrical and Information Engineering, Faculty of Engineering, Kiel University, Kiel, Germany, [2]Arbeitsgruppe Neurogeriatrie, Department of Neurology, Universitätsklinikum Schleswig-Holstein, Kiel, Germany, [3]Department of Biomedical Sciences, University of Sassari, Sassari, Italy, [4]Department of Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, [5]Laboratory of Movement Analysis and Measurement, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, [6]Department of Electronics and Telecommunications, Polytechnic of Turin, Turin, Italy, [7]Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom, [8]Gesellschaft für Medizinische Forschung, Robert-Bosch Foundation GmbH, Stuttgart, Germany, [9]INSIGNEO Institute for In Silico Medicine, The University of Sheffield, Sheffield, United Kingdom, [10]Department of Mechanical Engineering, The University of Sheffield, Sheffield, United Kingdom, [11]Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom, [12]School of Computing, Newcastle University, Newcastle upon Tyne, United Kingdom, [13]Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain, [14]Faculty of Health and Life Sciences, Universitat Pompeu Fabra, Barcelona, Spain, [15]CIBER Epidemiología y Salud Pública, Madrid, Spain, [16]Insight Centre for Data Analytics, University College Dublin, Dublin, Ireland, [17]School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland, [18]Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi", University of Bologna, Bologna, Italy, [19]Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRISDV), University of Bologna, Bologna, Italy, [20]Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom, [21]Grünenthal GmbH, Aachen, Germany, [22]Center for the Study of Movement, Cognition and Mobility, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, [23]Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel, [24]Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle upon Tyne, United Kingdom, [25]Novartis Institute of Biomedical Research, Novartis Pharma AG, Basel, Switzerland, [26]Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, United Kingdom, [27]Digital Health Department, CSEM SA, Neuchâtel, Switzerland, [28]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway

Introduction: The clinical assessment of mobility, and walking specifically, is still mainly based on functional tests that lack ecological validity. Thanks to inertial

measurement units (IMUs), gait analysis is shifting to unsupervised monitoring in naturalistic and unconstrained settings. However, the extraction of clinically relevant gait parameters from IMU data often depends on heuristics-based algorithms that rely on empirically determined thresholds. These were mainly validated on small cohorts in supervised settings.

Methods: Here, a deep learning (DL) algorithm was developed and validated for gait event detection in a heterogeneous population of different mobility-limiting disease cohorts and a cohort of healthy adults. Participants wore pressure insoles and IMUs on both feet for 2.5 h in their habitual environment. The raw accelerometer and gyroscope data from both feet were used as input to a deep convolutional neural network, while reference timings for gait events were based on the combined IMU and pressure insoles data.

Results and discussion: The results showed a high-detection performance for initial contacts (ICs) (recall: 98%, precision: 96%) and final contacts (FCs) (recall: 99%, precision: 94%) and a maximum median time error of −0.02 s for ICs and 0.03 s for FCs. Subsequently derived temporal gait parameters were in good agreement with a pressure insoles-based reference with a maximum mean difference of 0.07, −0.07, and <0.01 s for stance, swing, and stride time, respectively. Thus, the DL algorithm is considered successful in detecting gait events in ecologically valid environments across different mobility-limiting diseases.

KEYWORDS

deep learning (artificial intelligence), free-living, gait analysis, gait events detection, inertial measurement unit (IMU), mobility

1. Introduction

Mobility is the ability to move about in the home and community (1). Mobility can be affected by chronic health conditions, including but not limited to neurological, respiratory, cardiac, and musculoskeletal disorders (2). Deficits in mobility have been linked with a reduced quality of life, an increased fall risk, and mortality (2, 3), therefore, mobility is regarded as an essential aspect of health (4). The most common and functionally relevant aspect of mobility that is affected by aging and chronic health conditions is walking (1, 5).

To date, the clinical assessment of mobility is based on functional tests that include short walking tasks (6–9). A common shortcoming of these functional tests is the lack of ecological validity: Walking, as measured in clinical settings, does not reflect daily life walking (3, 10–12). The transition to unsupervised monitoring of human motion in naturalistic and unconstrained daily life activities is driven mainly using wearable inertial measurement units (IMUs) (4, 13). It is noteworthy that meanwhile both European and American notified bodies for the certification of medical devices (Medical Device Regulation and Food and Drug Administration, respectively) have put focus on wearable sensors by updating their regulations for the design, pre-clinical validation, and clinical validation of devices that include wearable IMUs (13, 14). Similarly, both the European Medicines Agency and the United States Food and Drug Administration encourage the inclusion of parameters from unsupervised patient monitoring as exploratory endpoints in clinical trials (11, 15).

A critical step for the objective analysis of gait is the segmentation of gait sequences into gait cycles (16–18), i.e., the basic repetitive unit that gait is comprised of (19, 20). The beginning and end of each gait cycle, also referred to as stride, are often determined from two successive initial contacts (ICs) of the same foot (19, 20). Together with the instant at which the foot leaves the ground (i.e., final contact, FC), each stride can be divided into a stance and swing phase (18–21). ICs and FCs are commonly referred to as gait events (19, 20, 22) and are a prerequisite for any further clinical gait analysis (18). The detection of ICs and FCs from IMUs is typically done using heuristics-based algorithms (23–30). Many of these algorithms use local maxima or minima of the acceleration and/or angular velocity signals along one axis (31), which requires knowledge of the sensor-to-segment alignment (32, 33). However, in unsupervised human gait monitoring, the sensor-to-segment alignment cannot be controlled as study participants often attach the sensor themselves, for example, after showering (34). Therefore, the technical validity of these algorithms for the case of unsupervised human gait monitoring is still an ongoing challenge also due to the scarcity of labeled free-living gait data (35–37). Additionally, IMU-based gait signals are affected by disease characteristics, participant activity levels, and the exact context in which walking takes please, and therefore, any heuristics-based algorithm that was developed based on lab-based gait data might not translate directly to free-living gait (3, 11, 15, 30, 38).

In contrast to the aforementioned heuristics-based algorithms, machine learning-based algorithms do not depend on user-defined sets of rules but rather learn to recognize gait signals directly from annotated data (39–41). Hidden Markov models (HMMs), for example, were successfully applied for gait segmentation in healthy (42, 43) and pathological gait (42, 44), but only in-lab recorded gait data were used to check for validity. A recent study

used HMMs to segment gait cycles from free-living gait data and reached 96% recall and 89% precision for free-living data, however, data were only from participants with Parkinson's disease (PD) (45). Although HMMs thus seem a good fit for modeling the sequential nature of the gait cycle, one still needs to define the number of discrete states beforehand, and it would be needed to have a separate model per activity if more than just gait was to be detected (46, 47). Deep learning (DL)-based algorithms provide an alternative approach that does not require any heuristic rules but rather learns relevant data representations automatically from a set of input features and reference annotations (40, 41, 48, 49). DL algorithms have been successfully applied for gait event detection from stereophotogrammetric data (50–54) and from inertial measurement unit data (34, 55), however, only for in-lab gait data.

Therefore, the specific aim of the current study was to determine whether a previously in-lab validated DL-based algorithm (34) for the detection of ICs and FCs can be used for the detection of gait events in pre-extracted real-life walking bouts in a heterogeneous cohort of different mobility-limiting diseases. For the current study, walking bouts were defined according to the recently published consensus framework for digital mobility monitoring (2).

2. Materials and methods

2.1. Data collection

2.1.1. Study participants

As part of the Mobilise-D technical validation study (56), a convenience sample of 108 participants was recruited at five independent study sites (Newcastle upon Tyne Hospitals NHS Foundation Trust, UK, Sheffield Teaching Hospitals NHS Foundation Trust, UK, Tel Aviv Sourasky Medical Center, Israel, Robert Bosch Foundation for Medical Research, Germany, University of Kiel, Germany). The sample represented five mobility-limiting disease cohorts [congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), multiple sclerosis (MS), Parkinson's disease (PD), and proximal femoral fracture (PFF)] and a cohort of healthy older adults (HA) (56). These cohorts cover a range of walking speed, mobility challenges, and potential events that are of clinical interest, such as improving vs. worsening of function, falls, hospitalization, nursing home admission, and death. Furthermore, as the participants were recruited at five different sites across Europe, they ensured a geographical representation and covered a diverse representation of healthcare organization, such as in- vs. outpatient care, as well as public vs. private health services (1, 56). Participants needed to be able to walk 4 m independently, to give informed consent, and have a Montreal Cognitive Assessment score $>$ 15 (57). A detailed description of inclusion and exclusion criteria is provided elsewhere (56), and ranges of values for cohort-specific clinical scales are detailed in Table 1.

2.1.2. Study protocol

Study participants were equipped with the INertial module with Distance sensors and Pressure insoles (INDIP) system that included both pressure insoles (PIs) and IMUs to record movement signals from both feet and the lower back (27, 58, 59). Participants wore the INDIP system for 2.5 h in their habitual environment, e.g., home, work, community, and/or outdoor environment, which was chosen by the participant, with no specific restrictions (56). To capture the largest possible range of activities, participants were provided with a list of activities that could be included if relevant to their chosen environment (e.g., rising from a chair, walking to another room, and walking outdoors). No supervision or structure as to how these tasks were completed was given to the participants. The duration of the observation has been established as a trade-off between experimental, clinical, and technical requirements (56).

2.2. Data processing

2.2.1. Data preparation

Data from the INDIP system were synchronized by setting the clock to have the same timestamp for all the sensors between the left and right foot, and values were recorded at a sampling frequency, f_s, of 100 Hz. As input to the DL algorithm, only the raw accelerometer and gyroscope data from both feet were used. Data were split into three different datasets: a training set, a validation set, and a testing set (40, 41). For this purpose, for each of the six cohorts, data from approximately 20% of the participants were assigned to the testing set, data from another 20% of the participants were assigned to the validation set, and data from the remaining participants were used as the training set.

The validation set was used to find an optimal network architecture using grid search (60), and the training set was used to optimize the corresponding model parameters (40, 41). The testing set was only used for the final evaluation, and notably, the numbers presented in the Section Results only corresponded to the performance of the testing set.

2.2.2. Reference system

For all data, the gait events, that is both ICs and FCs, were detected separately from the PIs and IMUs from the INDIP system that is described in detail elsewhere (61) to meet the emerging demands associated with reproducibility and replicability in biomedical research and regulatory qualification (62). Then, the results were combined, and priority was given to the PIs in case both modalities detected an event (63). For the PIs, foot-ground contact was defined when at least three sensing elements from the PI belonging to the same spatial neighborhood were consecutively activated and deactivated (64). For the IMUs, an existing algorithm, originally designed for shank-worn IMUs, was adapted for use with foot-worn IMUs. Previously, it was validated for the detection of supervised gait events in older, hemiparetic, parkinsonian, and choreic gait (27, 65) and across multiple research centers for parkinsonian and mildly cognitive impaired gait (66).

From these gait events, walking bouts (WBs) were formed by merging information from left and right strides (27, 28). Each WB represented a gait sequence with a minimum of two left and two right strides (2, 63). Here, strides were only considered valid if (i) the stride duration was between 0.2 and 3 s and (ii) the stride

TABLE 1 Dataset details for training, validation, and testing sets, including the total number of bouts and strides.

Set	Cohort	Number of participants	Age (years)	Height (cm)	Weight (kg)	Clinical scale (mean [min, max])	Number of bouts	Number of strides
Training	CHF	8	69 (13)	177 (8)	86 (20)	KCCQ: 81.8 [37.0, 96.3]	189	11326
	COPD	11	70 (9)	169 (6)	73 (14)	CAT: 21.1 [6.0, 33.0] FEV_1: 1.7 [0.9, 2.7]	187	6562
	MS	12	47 (8)	171 (14)	80 (23)	EDSS: 3.5 [1.0, 6.5]	139	6216
	PD	12	70 (7)	175 (6)	79 (16)	HandY: 2.0 [1.0, 3.0] UPDRS: 31.8 [6.0, 54.0]	165	7574
	PFF	10	83 (6)	172 (9)	71 (16)	SPPB: 7.3 [0.0, 12.0]	151	5838
	HA	12	71 (7)	168 (10)	76 (11)		245	13597
Validation	CHF	2	74 (13)	172 (21)	87 (3)	KCCQ: 94.8 [89.6, 100.0]	41	1210
	COPD	3	69 (14)	171 (10)	69 (12)	CAT: 15.3 [6.0, 26.0] FEV_1: 1.4 [1.3, 1.6]	68	1890
	MS	3	42 (15)	172 (13)	97 (24)	EDSS: 2.5 [1.5, 4.0]	24	863
	PD	3	70 (7)	174 (6)	79 (21)	HandY: 2.3 [2.0, 3.0] UPDRS: 28.0 [24.0, 33.0]	61	3466
	PFF	2	71 (1)	164 (8)	60 (9)	SPPB: 5.0 [1.0, 9.0]	31	1087
	HA	4	72 (4)	163 (10)	77 (18)		126	4952
Testing	CHF	2	65 (13)	168 (1)	77 (16)	KCCQ: 66.7 [47.9, 85.4]	10	407
	COPD	3	69 (8)	166 (3)	80 (18)	CAT: 18.7 [13.0, 24.0] FEV_1: 1.4 [0.8, 2.3]	79	2346
	MS	3	58 (12)	172 (16)	87 (25)	EDSS: 4.7 [3.0, 6.0]	53	2576
	PD	3	70 (11)	166 (11)	73 (8)	HandY: 2.3 [2.0, 3.0] UPDRS: 24.3 [7.0, 41.0]	38	2448
	PFF	2	76 (6)	168 (8)	75 (28)	SPPB: 6.5 [3.0, 10.0]	21	1649
	HA	4	73 (3)	164 (11)	72 (10)		94	3674

CAT, COPD assessment test; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; EDSS, Expanded disability status scale; FEV_1, Forced expiratory volume in 1 s; HA, healthy adults; HandY, Hoehn and Yahr scale; KCCQ, Kansas City cardiomyopathy questionnaire; MS, multiple sclerosis; PD, Parkinson's disease; PFF, proximal femoral fracture; SPPB, short physical performance battery; UPDRS, Movement Disorder Society-sponsored Unified Parkinson's Disease Rating Scale, part III. Age, height, and weight are presented as mean (standard deviation), and the clinical scales are presented as mean [minimum, maximum].

length was minimally 0.15 m. A resting period of 3 s determined consecutive WBs, thus, each WB could contain a resting period of $\leq$3 s.

For the current study, we analyzed only those WBs that lasted $\geq$10 s (67–70) and for which both the INDIP's PIs and IMUs were used for determining the gait events. These gait events were considered as reference annotations for training and evaluating the DL algorithm.

2.2.3. Deep learning algorithm

The DL algorithm was based on the neural network (NN) that was previously validated on in-lab gait data from shank-worn IMUs worn by participants with different neurological diseases (34, 71). At the core of the NN was a temporal convolutional network (TCN) (72, 73). The TCN was built from stacking residual blocks (74) with an exponentially increasing dilation factor for the convolutional layers (Figure 1).

Specifically, each residual block comprised two sequences of a dilated convolution (Conv) layer (75), a batch normalization (BatchNorm) layer (76), a rectified linear unit (ReLU) activation layer, and a dropout layer (77). A residual connection was used to perform convolution with a kernel size of 1 in case the number of feature maps did not match the number of input channels (72, 73). The outputs of the second dropout layer and the residual connection were summed elementwise and inputted to a ReLU activation layer. The convolution layers consisted of 64 filters with a kernel size of 3 and a dilation factor of 2^{m-1} with $m = 1, \cdots, N_{dil}$ for the m-th residual block (with $N_{dil} = 6$, the number of residual blocks, and thus, the maximum dilation factor was $2^5 = 32$).

The outputs of the last residual block were passed through a fully connected (also referred to as dense) layer followed by a softmax activation layer (78, 79). The final outputs were then regarded as probability that a certain gait event took place at the given time step, t_n.

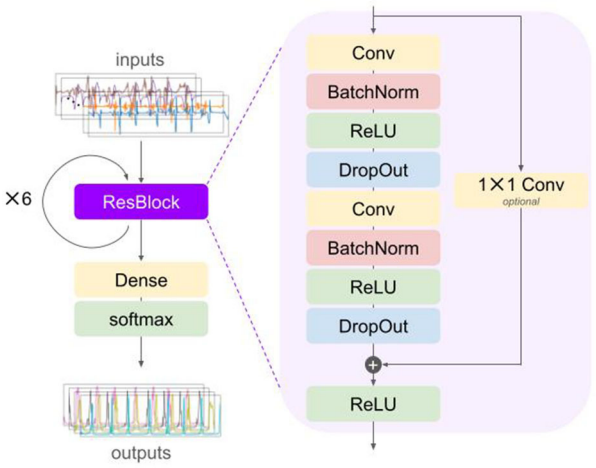

FIGURE 1
Schematic depiction of the deep learning model architecture with a residual block (ResBlock) that is repeated (in this case, six times) before a dense and softmax layer are applied. Inputs to the network are the raw accelerometer and gyroscope data of both left and right inertial measurement units. The outputs are estimated probabilities for each of the gait events for each time step. BatchNorm, batch normalization; Conv, convolution; DropOut, dropout; ReLU, recitified linear unit.

2.3. Evaluation

As in our previous studies (34), the performance was evaluated with the testing set only. The trained model was used to predict the probability that any gait event occurred from the IMU data. Peak probabilities, with a minimum probability, $\Delta_{\mathrm{Pr}} = 0.5$, and a minimum interpeak distance, $\Delta_t = 0.5\,\mathrm{s}$, were considered detected events.

Performance was evaluated for the overall detection performance, time agreement between predicted and annotated gait event timings, and time agreement between subsequently derived stride-specific gait parameters.

2.3.1. Overall detection performance

The overall detection performance quantified how many of the annotated gait events were detected (true positives), how many of the annotated gait events were not detected (false negatives), and how many of the detected events were not annotated (false positives). From these numbers, the recall (also referred to as sensitivity) and precision (also referred to as positive predictive value) were calculated as follows:

$$recall = \frac{\#\ true\ positives}{\#\ true\ positives + \#\ false\ negatives}, \quad (1)$$

$$precision = \frac{\#\ true\ positives}{\#\ true\ positives + \#\ false\ positives}. \quad (2)$$

Thus, the recall represented the fraction of annotated events that were detected, and the precision represented the fraction of events that were truly gait events.

Here, in case the absolute time difference between an annotated and predicted event was ≤250 ms, it was considered a true positive event (30, 34, 80, 81) (in other words, a tolerance window of 500 ms centered around the reference timing was used).

2.3.2. Time agreement

For all correctly detected gait events (true positives), the time agreement between the detected and annotated event timings was quantified by

$$\epsilon = t_{ref} - t_{pred}, \quad (3)$$

where t_{pred} is the timing corresponding to the peak probability and t_{ref} is the timing of the INDIP-derived annotations.

As a robust measure for the time agreement and its spread, the median time error and the inter-quartile range (IQR) were computed (82), and time agreements were visualized using box plots.

2.3.3. Stride-specific gait parameters

For those strides where both ICs and the FC in between were detected, the stance, swing, and stride times were computed (19, 20, 83). Stance time was the time between an FC and the preceding IC of the same foot, swing time was the time between an IC and the preceding FC of the same foot, and stride time was the time between two consecutive ICs of the same foot (34, 83).

For each of these temporal gait parameters, the mean time difference and the limits of agreement (LoA) based on a 95% confidence interval (CI) were computed (82). Differences were visualized using Bland–Altman plots (84, 85).

3. Results

3.1. Demographics

Data were collected from 108 different participants, and eventually data from 99 participants were used for the current study (Table 1). Data from the other participants were excluded due to incomplete or missing data from the INDIP system or because no WBs ≥ 10 s were recorded. Eventually, the DL-based algorithm was evaluated for its performance in detecting gait events of 13,100 strides divided over 295 bouts recorded from 17 participants in the testing set.

3.2. Overall detection performance

The overall detection performance was quantified by the number of true positives, number of false negatives, and number of false positives. From these numbers, the recall and precision were calculated (Table 2). In total, from 13,134 ICs, the algorithm detected 12,985 events (i.e., 99%) and missed 169 events (i.e., 1%), and similarly, from 12,838 FCs, the algorithm detected 12,747 events (i.e., 99%) and missed 91 events (i.e., 1%). When evaluated per cohort, the recall for the IC detection was ≥98%, and the precision was ≥96%. Similarly, the recall was ≥99%, and the precision was ≥94% for FC detection for all cohorts.

10.3389/fneur.2023.1247532

TABLE 2 Overall detection performance of initial and final contacts evaluated per cohort.

Cohort	Initial contacts						Final contacts					
	TP	FN	FP	R (%)	P (%)	F1 (%)	TP	FN	FP	R (%)	P (%)	F1 (%)
CHF	408	3	18	99	96	97	401	1	23	100	95	97
COPD	2,294	58	86	98	96	97	2,235	37	147	99	94	96
MS	2,563	19	72	99	97	98	2,518	12	95	100	96	98
PD	2,431	23	40	99	98	99	2,411	3	55	100	98	99
PFF	1,642	11	15	99	99	99	1,614	18	45	99	97	98
HA	3,627	55	98	99	97	98	3,568	20	141	99	96	98

CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; F1, F1 score; FN, false negative; FP, false positive; HA, healthy adults; MS, multiple sclerosis; P, precision; PD, Parkinson's disease; PFF, proximal femoral fracture; R, recall; TP, true positive.

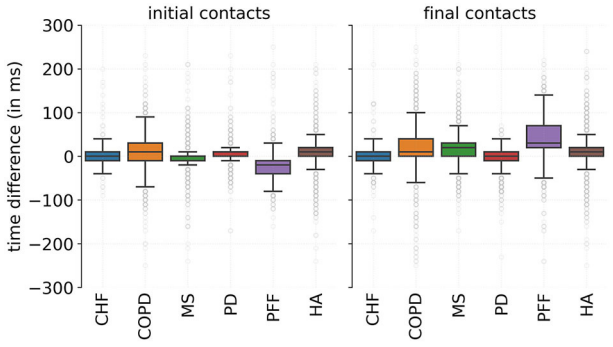

FIGURE 2
Time difference between the predicted and reference events timings for initial and final contacts evaluated per cohort. A positive time difference corresponded to an advanced detection. CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; MS, multiple sclerosis; HA, healthy adults; PD, Parkinson's disease; PFF, proximal femoral fracture.

3.3. Time agreement

For all the correctly detected events, i.e., true positives, the difference between the detected event timing and the annotated event timings was calculated according to Equation (6). The median time error was close to 0 s with the IQR enclosing a zero difference for both ICs and FCs for all cohorts, except for the PFF cohort (Figure 2). The PFF cohort showed a median time error of −0.02 s and an IQR of 0.03 s for IC detection, and a median time error of 0.03 s and IQR of 0.05 s for FC detection (Table 3).

3.4. Stride-specific gait parameters

For those strides that had two correctly detected ICs and a correctly detected FC in between, stride-specific temporal gait parameters (i.e., stance time, swing time, and stride time) were calculated. For all cohorts, the mean differences between the stance, swing, and stride times derived from the detected events and those derived from the annotations were close to zero with the LoA encapsulating a zero-mean difference (Figure 3). Notably, for the PFF cohort, the mean time difference for the stance time was +0.07 s, and the mean time difference for the swing time was −0.07 s, which resulted in a zero-mean difference for the stride time (Table 4). Similarly, for all gait phases, the absolute errors were 0.04 s or less for all cohorts, except the PFF cohort (Table 5). This resulted in a relative time error for the stride times of ≤2% across all cohorts, but for the swing times, the relative time error for the PFF cohort was 27%, and for the COPD cohort, it was 12%.

4. Discussion

The specific aim of the current study was to determine whether a previously in-lab validated DL-based gait event detection algorithm (34) could be used for the detection of gait events from real-life walking bouts in a heterogeneous cohort of different

TABLE 3 Time differences between the predicted event timings and the annotated event timings evaluated per cohort.

Cohort	Initial contacts		Final contacts	
	Median (ms)	IQR (ms)	Median (ms)	IQR (ms)
CHF	0	20	0	20
COPD	10	40	10	40
MS	0	10	20	30
PD	10	10	20	30
PFF	−20	30	30	50
HA	10	20	10	20

CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; MS, multiple sclerosis; HA, healthy adults; IQR, inter-quartile range; PD, Parkinson's disease; PFF, proximal femoral fracture.

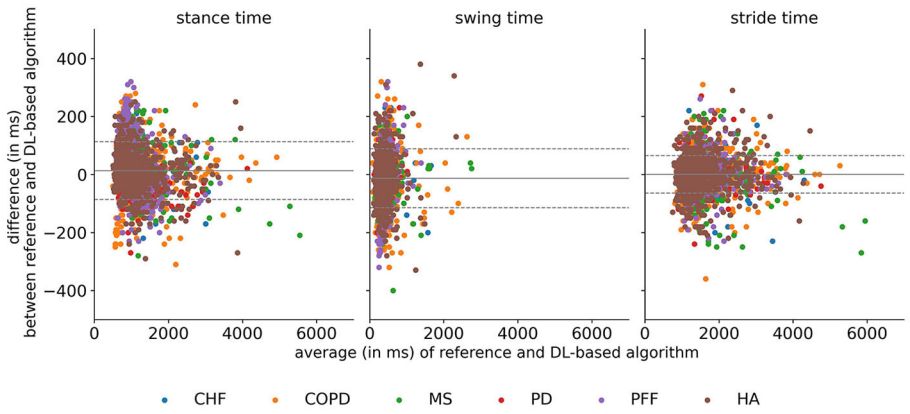

FIGURE 3
Bland–Altman plots for the stance, swing, and stride times evaluated per cohort. The gray solid line corresponds to the overall mean difference, and the dashed lines correspond to the mean difference ± 1 standard deviation. CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; DL, deep learning; HA, healthy adults; MS, multiple sclerosis; PD, Parkinson's disease; PFF, proximal femoral fracture.

TABLE 4 Mean differences (bias) and limits of agreement for a 95% confidence interval for the stance, swing, and strides evaluated for each cohort.

Cohort	Stance time		Swing time		Stride time	
	Mean difference (s)	LoA (s, s)	Mean difference (s)	LoA (s, s)	Mean difference (s)	LoA (s, s)
CHF	−0.00	(−0.08, 0.07)	0.00	(−0.07, 0.07)	−0.00	(−0.07, 0.07)
COPD	0.01	(−0.11, 0.13)	−0.01	(−0.13, 0.11)	0.00	(−0.08, 0.08)
MS	0.02	(−0.05, 0.10)	−0.02	(−0.10, 0.05)	−0.00	(−0.06, 0.06)
PD	−0.01	(−0.06, 0.04)	0.01	(−0.04, 0.06)	0.00	(−0.05, 0.05)
PFF	0.07	(−0.06, 0.19)	−0.07	(−0.20, 0.07)	0.00	(−0.07, 0.07)
HA	0.00	(−0.07, 0.08)	−0.00	(−0.09, 0.08)	0.00	(−0.07, 0.07)

CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; HA, healthy adults; LoA, limits of agreement; MS, multiple sclerosis; PD, Parkinson's disease; PFF, proximal femoral fracture.

mobility-limiting diseases. For that purpose, participants from different disease cohorts (CHF, COPD, MS, PD, and PFF) and a cohort of healthy adults were equipped with the INDIP system that consisted of PIs and IMUs for both feet. Participants wore the INDIP system for 2.5 h in the habitual environment, as chosen by the participants, and a wide range of activities were recorded in these ecologically valid environments. Data from the PIs and IMUs were used to generate reference timings for ICs and FCs, whereas raw data from the accelerometer and gyroscope were used as the input to the DL algorithm to identify ICs and FCs.

The recall and precision of gait events were used as a general measure for the detection performance and were considered high (i.e., recall $\geq$ 98% and precision $\geq$ 96%). For comparison, in Trojaniello et al. (27), no missed or extra gait events were observed in a heterogeneous sample of elderly, hemiparetic, parkinsonian, and choreic gait, but data were only collected from walking back

TABLE 5 Stance, swing, and stride times obtained from the reference and the DL algorithm, and the absolute and relative time errors for comparison.

Gait phase	Cohort	Reference system		DL algorithm		Absolute error		Relative error	
		s	(s, s)	s	(s, s)	S	(s, s)	%	(%, %)
Stance	CHF	0.93	(0.90, 0.97)	0.94	(0.91, 0.97)	0.03	(0.02, 0.03)	3	(2, 3)
	COPD	0.93	(0.91, 0.94)	0.92	(0.90, 0.93)	0.04	(0.04, 0.05)	5	(5, 5)
	MS	0.98	(0.97, 0.99)	0.96	(0.94, 0.97)	0.03	(0.03, 0.03)	3	(3, 3)
	PD	0.80	(0.79, 0.80)	0.81	(0.80, 0.81)	0.02	(0.02, 0.02)	2	(2, 2)
	PFF	0.90	(0.88, 0.91)	0.83	(0.82, 0.84)	0.08	(0.08, 0.08)	9	(9, 9)
	HA	0.84	(0.83, 0.85)	0.84	(0.83, 0.85)	0.03	(0.02, 0.03)	3	(3, 3)
Swing	CHF	0.41	(0.40, 0.42)	0.41	(0.39, 0.42)	0.02	(0.02, 0.03)	6	(5, 7)
	COPD	0.43	(0.42, 0.43)	0.43	(0.43, 0.44)	0.04	(0.04, 0.05)	12	(11, 13)
	MS	0.41	(0.41, 0.42)	0.44	(0.43, 0.44)	0.03	(0.03, 0.03)	9	(8, 9)
	PD	0.41	(0.40, 0.41)	0.40	(0.39, 0.40)	0.02	(0.02, 0.02)	4	(4, 4)
	PFF	0.34	(0.34, 0.35)	0.41	(0.40, 0.41)	0.08	(0.08, 0.08)	27	(26, 28)
	HA	0.36	(0.36, 0.36)	0.36	(0.36, 0.37)	0.03	(0.03, 0.03)	8	(8, 9)
Stance	CHF	1.34	(1.31, 1.38)	1.34	(1.31, 1.38)	0.02	(0.02, 0.02)	1	(1, 2)
	COPD	1.35	(1.33, 1.37)	1.35	(1.33, 1.37)	0.02	(0.02, 0.02)	2	(2, 2)
	MS	1.39	(1.38, 1.40)	1.39	(1.38, 1.40)	0.02	(0.02, 0.02)	1	(1, 1)
	PD	1.20	(1.19, 1.21)	1.20	(1.19, 1.21)	0.01	(0.01, 0.01)	1	(1, 1)
	PFF	1.24	(1.22, 1.25)	1.24	(1.22, 1.25)	0.02	(0.02, 0.02)	2	(2, 2)
	HA	1.20	(1.19, 1.21)	1.20	(1.19, 1.21)	0.02	(0.02, 0.02)	2	(1, 2)

Values represent the mean and 95% confidence interval of all stances, swings, and strides of the test subjects for the given cohort. CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; DL, deep learning; HA, healthy adults; MS, multiple sclerosis; PD, Parkinson's disease; PFF, proximal femoral fracture.

and forth for 1 min in a 12 m walkway. Similarly, high recall and precision (≥98%) were reported for a continuous wavelet transform (CWT)-based algorithm, but it was evaluated only for 13 healthy participants and 3 hemiplegic participants who walked continuously along a 10 m walkway (86). A recent study (45) found a recall of 96% and precision of 89% in a cohort of 28 PD participants, who wore two IMUs on the feet for 2 weeks, which are slightly lower than the recall and precision from the current study. Overall, the data of the studies presented here, including the present study, indicate that very high recall and precision values can be achieved with the deep learning approach for the detection of gait events. This, together with the higher flexibility of the DL-based algorithms compared to conventional algorithms, speaks for the future use of such algorithms for the detection of gait in mobility-limiting diseases also in the habitual environment.

For the correctly detected gait events, the time differences between the predicted event timing and the annotated event timings were quantified as a measure of temporal agreement between the reference system and the DL-based algorithm. The time differences were still in the same range as those previously reported for CWT-based (23, 27, 30, 86, 87) and DL-based algorithms (34) validated on in-lab gait data. To put this into perspective, studies that evaluated the time differences of detected gait events from PIs when compared to force plates or instrumented walkways also reported time differences in the range from 0.02 s to 0.04 s (17, 64, 87). For the INDIP pressure insole method, a negligible delay (<10 ms) was observed for FCs, and a consistent IC anticipation (20 ms) was found when compared to force plates (64). It suggests that a certain margin of uncertainty should be considered when interpreting gait event timing differences in the DL-based algorithm.

Finally, stride-specific gait parameters were derived for the correctly detected events. These may be of greatest clinical relevance since changes in spatiotemporal gait parameters were associated with a shorter time to PD diagnosis (88) and from mild cognitive impairment to Alzheimer's disease (89), and values of temporal gait parameters were different in disease cohorts compared to healthy cohorts (90–92). Here, a zero-mean time difference was found for the stride times for all cohorts. Similarly, the time differences for stance and swing times were centered around a zero-mean difference for all cohorts, with only the mean differences for the stance and swing time of the PFF cohort being a bit larger (0.07 s and −0.07 s for the stance and swing time, respectively). The mean differences for stance and swing times in the PFF cohort may in part be explained by the altered gait pattern that is observed in this cohort (93, 94). Nonetheless, the time agreement for the stride-specific temporal gait parameters derived from the DL algorithm and the reference system was in a similar range as those communicated before for a comparable DL-based approach that evaluated results only from straight-line walking in a supervised laboratory setting (55).

The very good results that were obtained in the current study for two-feet-worn IMUs (56) combined with the results for a single shank-worn IMU from our previous study (34) provided evidence that the algorithm performance generalizes to other sensor wear locations and to free-living gait data. The current algorithm has the additional benefit that it does not require the knowledge of exact sensor location and orientation relative to the feet contrary to many previously validated algorithms (23, 24, 31, 34, 95). This has the practical consequence that there are less stringent requirements for study participants or future patients on how to attach the sensors to their feet. Since for the previous validation the input data consisted of the raw accelerometer and gyroscope signals from a single sensor that was located either laterally above the ankle joint or medially below the knee joint (34), the algorithm for the current validation was again trained, validated, and tested. Both studies show a high recall and sensitivity, highlighting that the algorithm is capable of detecting gait events from different sensor locations without the loss of accuracy provided that sufficient training data are available for any new sensor location (34). Furthermore, the algorithm performance was evaluated across a broad spectrum of five different mobility-limiting disease cohorts, and although the number of participants in the testing set for each cohort was low, it showed that the algorithm was able to accurately detect gait events in heterogeneous pathological gait patterns. This will ultimately allow future users of the algorithm to perform not only sensitivity analyses for individual cohorts but also specificity analyses across different cohorts.

The limitation of the current study included that only data from detected WBs were used. This means that gait event detection relied on the accurate detection of gait sequences as a preceding step (45). However, several algorithms have been reported for accurate IMU-based gait sequence detection in both healthy and disease cohorts (24, 25, 28, 29, 46, 96–100). Furthermore, data from some participants had to be excluded from analysis due to missing or incomplete data which was mainly due to issues with the PIs. As reference timings for gait events are still obtained mainly from force or pressure measuring device (23), it showed the difficulty of obtaining a dataset with annotated gait events on completely unsupervised free-living gait data (35–37, 45). To get a better picture of the algorithm's generalizability to other datasets, it needs to be tested on newly unseen datasets, for example, with a slightly different sensor setup, such as in Martindale et al. (46).

In addition, the study did not evaluate clinical aspects in detail, such as medication and symptom fluctuations. This is, in part, due to the heterogeneous sample of participants with different mobility-limiting diseases. Consequently, the current study did not focus on identifying, for example, digital biomarkers of disease progression, for which a greater sample size of a specific disease would be required. However, as this is a study comparing, in the same person, systems at one point in time on a motility aspect, we believe that this does not influence the results reported here. Furthermore, it should be stressed that the heterogeneous sample is an asset of the current study as the results show that the algorithm achieves excellent performance for different pathological gait patterns. Given the time span of 2.5 h, we did not specifically investigate whether disease-associated gait abnormalities, such as freezing of gait in PD (101), were captured by the recording. However, the duration of the assessment was chosen as a trade-off between experimental, clinical, and technical requirements (56) and is five times longer than the recommendations for validation procedures of assessing physical activity in older adults (102). Lastly, the current analysis also relied on a peak detection algorithm to identify the most probable timings of gait events (34, 46, 55). However, from a clinical perspective, this may be regarded as a benefit since it would allow a clinician to decide whether to consider certain strides based on how confidently it can be assumed that it was indeed a stride.

5. Conclusion

This study aimed to validate a DL algorithm for the detection of gait events in an ecologically valid environment across different mobility-limiting disease cohorts. The performance evaluation showed an excellent detection rate and low time errors for both event timings and subsequently derived temporal gait parameters for all cohorts. The DL reached a performance that was in a similar range or slightly better than approaches that were to date only validated on in-lab recorded gait data or for a specific disease cohort.

As the DL algorithm does not rely on expert-defined decision rules or hand-crafted features nor on exact sensor-to-segment alignment, it poses fewer requirements on the data collection.

Our next steps include extending the current analysis for data from multiple days and evaluating to which extent the DL network can be trained using participant-specific data to improve gait event detection on an individual level. Future studies may also consider the development of novel gold-standard systems that allow validation approaches beyond lower limb movement, for example, to include upper limb movement.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The Mobilise-D consortium are planning to make representative data available through online repositories (Zenodo & GitHub) in the near future. All scripts that were designed for the current analysis are available online and can be found at: GitHub, https://github.com/neurogeriatricskiel/mobilised. Requests to access these datasets should be directed to RR, r.romijnders@neurologie.uni-kiel.de.

Ethics statement

The studies involving humans were approved by (1) Ethical Committee of the Medical Faculty of Kiel University, D438/18; (2) Helsinki Committee, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, 0551-392 19TLV; (3) Ethical Committee of the Medical Faculty of The University of Tuebingen, 647/2019BO2; and (4) London—Bloomsbury Research Ethics Committee, 19/LO/1507. The studies were conducted in accordance with the local legislation and institutional requirements. Written

informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

RR, FS, CH, ACe, GS, and WM: conceptualization and project administration. RR, FS, ACe, and GS: methodology. RR and FS: software. RR: validation, formal analysis, writing—original draft preparation, visualization, and investigation. ACe and WM: resources and funding acquisition. RR, FS, and ACe: data curation. FS, CH, ACe, GS, and WM: writing—reviewing and editing. GS and WM: supervision. RR, FS, CH, AKü, AP-I, ACe, LA, KA, CB, SB, TB, PB, EB, ACa, A-EC, MC, BC, LC, ID'A, SD, BE, SF, MF, JG, EG, JH, HH, EH, AKe, CK, FK, SK, CM, DM, EM-A, AM, LP, LR, LS, KS, BS, DS, AS, MU, BV, IV, AY, GS, and WM: intellectual contribution. All authors have read and agreed to the published version of the manuscript.

Funding

The authors acknowledge financial support by Land Schleswig-Holstein within the funding program Open Access Publikationsfonds. This study was supported by the Mobilise-D project that has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 820820. This JU receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA). LA, LR, AY, and SD were also supported by the National Institute for Health Research (NIHR) Newcastle Biomedical Research Centre (BRC) based at Newcastle upon Tyne Hospital NHS Foundation Trust and Newcastle University and the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon the Tyne Hospitals NHS Foundation Trust. This study was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) through the project B9 of the Collaborative Research Centre CRC 1261 Magnetoelectric Sensors: From Composite Materials to Biomagnetic Diagnostics.

Acknowledgments

Foremost, the authors would like to thank all study participants for their time and effort. The authors would also like to acknowledge all fellow members of the Mobilise-D WP2 for their input and discussions.

Conflict of interest

AM and FK are employees of, and may hold stock in, Novartis. BE reports consulting activities with, Siemens AG, Siemens Healthineers AG, WS Audiology GmbH outside of the study. He is a shareholder in Portabiles HealthCare Technologies GmbH. In addition, BE holds a patent related to gait assessment. MF is an employee of Grünenthal. LP and LC are co-founders and own shares of mHealth Technologies (https://mhealthtechnologies.it/). LS and CB are consultants of Philips Healthcare, Bosch Healthcare, Eli Lilly, and Gait Up. JH reports having submitted a patent for assessment of mobility using wearable sensors in Parkinson's disease. The intellectual property rights are held by the Tel Aviv Medical Center.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The content in this publication reflects the authors' view, and neither IMI nor the European Union, EFPIA, NHS, NIHR, DHSC, or any associated partners are responsible for any use that may be made of the information contained herein.

References

1. Rochester L, Mazzà C, Mueller A, Caulfield B, McCarthy M, Becker C, et al. A roadmap to inform development, validation and approval of digital mobility outcomes: the Mobilise-D approach. *Digit Biomark.* (2020) 4:13–27. doi: 10.1159/000512513

2. Kluge F, Del Din S, Cereatti A, Gaßner H, Hansen C, Helbostad JL, et al. Consensus based framework for digital mobility monitoring. *PLoS ONE.* (2021) 16:e0256541. doi: 10.1371/journal.pone.0256541

3. Warmerdam E, Hausdorff JM, Atrsaei A, Zhou Y, Mirelman A, Aminian K, et al. Long-term unsupervised mobility assessment in movement disorders. *Lancet Neurol.* (2020) 19:462–70. doi: 10.1016/S1474-4422(19)30397-7

4. Polhemus A, Delgado Ortiz L, Brittain G, Chynkiamis N, Salis F, Gaßner H, et al. Walking on common ground: a cross-disciplinary scoping review on the clinical utility of digital mobility outcomes. *NPJ Digit Med.* (2021) 4:149. doi: 10.1038/s41746-021-00513-5

5. Snijders AH, van de Warrenburg BP, Giladi G, Bloem BR. Neurological gait disorders in elderly people: clinical approach and classification. *Lancet Neurol.* (2007) 6:63–74. doi: 10.1016/S1474-4422(06)70678-0

6. Butland RJ, Pang J, Gross ER, Woodcock AA, Geddes DM. Two-, six-, and 12-minute walking tests in respiratory disease. *Brit Med J.* (1982) 284:1607–8. doi: 10.1136/bmj.284.6329.1607

7. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol.* (1994) 49:M85–94. doi: 10.1093/geronj/49.2.M85

8. Motl RW, Cohen JA, Benedict R, Phillips G, LaRocca N, Hudson LD, et al. Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis. *Mult Scler.* (2017) 23:704–10. doi: 10.1177/1352458517690823

9. Podsiadlo D, Richardson S. The timed 'up and go': a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc.* (1991) 39:142–8. doi: 10.1111/j.1532-5415.1991.tb01616.x

10. Atrsaei A, Corrà MF, Dadashi D, Vila-Chã N, Maia L, Mariani B, et al. Gait speed in clinical and daily living assessments in parkinson's disease patients: performance versus capacity. *NPJ Parkinsons Dis.* (2021) 7:24. doi: 10.1038/s41531-021-00171-0

11. Fasano A, Mancini M. Wearable-based mobility monitoring: the long road ahead. *Lancet Neurol.* (2020) 19:378–9. doi: 10.1016/S1474-4422(20)30033-8

12. Hillel I, Gazit E, Nieuwboer A, Avanzino L, Rochester L, Cereatti A, et al. Is every-day walking in older adults more analogous to dual-task walking or to usual walking? elucidating the gaps between gait performance in the lab and during 24/7 monitoring. *Eur Rev Aging Phys Act.* (2019) 16:6. doi: 10.1186/s11556-019-0214-5

13. Picerno P, Iosa M, D'Souza C, Benedetti MG, Paolucci S, Morone G. Wearable inertial sensors for human movement analysis: a five-year update. *Expert Rev Med Dev.* (2021) 18:79–94. doi: 10.1080/17434440.2021.1988849

14. Ravizza A, De Maria C, Di Pietro L, Sternini F, Audenino AL, Bignardi C, et al. Comprehensive review on current and future regulatory requirements on wearable sensors in preclinical and clinical testing. *Front Bioeng Biotechnol.* 7:313. doi: 10.3389/fbioe.2019.00313

15. Coran P, Goldsack JC, Grandinetti CA, Bakker JP, Bolognese M, Dorsey ER, et al. Advancing the use of mobile technologies in clinical trials: recommendations from the clinical trials transformation initiative. *Digit Biomark.* (2019) 3:145–54. doi: 10.1159/000503957

16. Celik Y, Stuart S, Woo WL, Godfrey A. Gait analysis in neurological populations: progression in the use of wearables. *Med Eng Phys.* (2021) 87:9–29. doi: 10.1016/j.medengphy.2020.11.005

17. Bobić VN, Djurić-Jovièić MD, Radovanović SM, Dragaević NT, Kostić VS, Popović MB. Challenges of stride segmentation and their implementation for impaired gait. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* (2018). p. 2284–7. doi: 10.1109/EMBC.2018.8512836

18. Rueterbories J, Spaich EG, Larsen B, Andersen OK. Methods for gait event detection and analysis in ambulatory systems. *Med Eng Phys.* (2010) 32:545–52. doi: 10.1016/j.medengphy.2010.03.007

19. Perry J, Burnfield JM. *Gait Analysis: Normal and Pathological Function.* 2nd ed. Thorofare, NJ: SLACK Incorporated (2010).

20. Richards J, Levine D, Whittle M. *Whittle's Gait Analysis.* 5th ed. Edinburgh: Churchill Livingstone (2012).

21. Bruening DA, Ridge ST. Automated event detection algorithms in pathological gait. *Gait Post.* (2014) 39:472–7. doi: 10.1016/j.gaitpost.2013.08.023

22. Hennerici M, Bäzner H. *Gangstörungen*. Berlin, Heidelberg: Springer (2001). doi: 10.1007/978-3-642-59529-5

23. Aminian K, Najafi B, Bula C, Leyvraz P-F, Robert PH. Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *J Biomech.* (2002) 35:689–99. doi: 10.1016/S0021-9290(02)00008-8

24. Salarian A, Russmann H, Vingerhoets FJG, Dehollain C, Blanc Y, Burkhard PR, et al. Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring. *IEEE Trans Biomed Eng.* (2004) 51:1434–43. doi: 10.1109/TBME.2004.827933

25. Paraschiv-Ionescu A, Buchser EE, Rutschmann B, Najafi B, Aminian K. Ambulatory system for the quantitative and qualitative analysis of gait and posture in chronic pain patients treated with spinal cord stimulation. *Gait Post.* (2004) 20:113–25. doi: 10.1016/j.gaitpost.2003.07.005

26. Trojaniello D, Cereatti A, Croce UD. Gait direction of progression changes using shank worn MIMUs. Validation on healthy and TBI subjects. In: *IV National Congress of National Group of Bioengineering (GNB).* Pavia (2014).

27. Trojaniello D, Cereatti A, Pelosin E, Avanzino L, Mirelman A, Hausdorff JM, et al. Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait. *J Neuroeng Rehabil.* (2014) 11:152. doi: 10.1186/1743-0003-11-152

28. Paraschiv-Ionescu A, Newman CJ, Carcreff L, Gerber CN, Armand S, Aminian K. Locomotion and cadence detection using a single trunk-fixed accelerometer: validity for children with cerebral palsy in daily life-like conditions. *J Neuroeng Rehabil.* (2019) 16:24. doi: 10.1186/s12984-019-0494-z

29. Paraschiv-Ionescu A, Soltani A, Aminian K. Real-world speed estimation using single trunk IMU: methodological challenges for impaired gait patterns. In: *2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* (2020). p. 4596–9. doi: 10.1109/EMBC44109.2020.9176281

30. Romijnders R, Warmerdam E, Hansen C, Welzel J, Schmidt G, Maetzler W. Validation of IMU-based gait event detection during curved walking and turning in older adults and Parkinson's disease patients. *J Neuroeng Rehabil.* (2021) 18:28. doi: 10.1186/s12984-021-00828-0

31. Niswander W, Kontson K. Evaluating the impact of IMU sensor location and walking task on accuracy of gait event detection algorithms. *Sensors.* (2021) 21:3989. doi: 10.3390/s21123989

32. Pacher L, Chatellier C, Vauzelle R, Fradet L. Sensor-to-segment calibration methodologies for lower-body kinematic analysis with inertial sensors: a systematic review. *Sensors.* (2020) 20:3322. doi: 10.3390/s20113322

33. Leineweber MJ, Gomez Orozco MD, Andrysek J. Evaluating the feasibility of two post-hoc correction techniques for mitigating posture-induced measurement errors associated with wearable motion capture. *Med Eng Phys.* (2019) 71:38–44. doi: 10.1016/j.medengphy.2019.06.013

34. Romijnders R, Warmerdam E, Hansen C, Schmidt G, Maetzler W. A deep learning approach for gait event detection from a single Shank-Worn IMU: validation in healthy and neurological cohorts. *Sensors.* (2022) 22:3859. doi: 10.3390/s22103859

35. Camps J, Samà A, Martín M, Rodríguez-Martín D, Pérez-López C, Moreno Arostegui JM, et al. Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit. *Knowledge Based Syst.* (2018) 139:119–31. doi: 10.1016/j.knosys.2017.10.017

36. Hammerla NY, Halloran S, Plötz T. Deep, convolutional, and recurrent models for human activity recognition using wearables. In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*. New York, NY (2016). p. 1533–40.

37. Raykov YP, Evers LJW, Badawy R, Bloem BR, Heskes TM, Meinders MJ, et al. Probabilistic modelling of gait for robust passive monitoring in daily life. *IEEE J Biomed Health Inform.* (2021) 25:2293–304. doi: 10.1109/JBHI.2020.3037857

38. Pacini Panebianco G, Cristina Bisi M, Stagni R, Fantozzi S. Analysis of the performance of 17 algorithms from a systematic review: influence of sensor position, analysed variable and computational approach in gait timing estimation from IMU measurements. *Gait Post.* (2018) 66:76–82. doi: 10.1016/j.gaitpost.2018.08.025

39. Haji Ghassemi N, Hannink J, Martindale CF, Gaßner H, Müller M, Klucken J, et al. Segmentation of gait sequences in sensor-based movement analysis: a comparison of methods in Parkinson's disease. *Sensors.* (2018) 18:145. doi: 10.3390/s18010145

40. Géron A. *Hands-On Machine Learning*. 2nd ed. Sebastopol, CA: O'Reilly Media, Inc. (2019).

41. Chollet F. *Deep Learning with Python*. 2nd ed. Shelter Island, NY: Manning Publications (2021).

42. Mannini A, Sabatini AM. A hidden Markov model-based technique for gait segmentation using a foot-mounted gyroscope. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. (2011). p. 4369−73. doi: 10.1109/IEMBS.2011.6091084

43. Panahandeh G, Mohammadiha N, Leijon A, Händel P. Continuous hidden Markov model for pedestrian activity classification and gait analysis. *IEEE Trans Instrument Meas.* (2013) 62:1073–83. doi: 10.1109/TIM.2012.2236792

44. Martindale CF, Strauss M, Gassner H, List J, Muller M, Klucken J, et al. Segmentation of gait sequences using inertial sensor data in hereditary spastic paraplegia. In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. p. 1266–69. Seogwipo: IEEE (2017). doi: 10.1109/EMBC.2017.8037062

45. Roth N, Küderle A, Ullrich M, Gladow T, Marxreiter F, Klucken J, et al. Hidden Markov model based stride segmentation on unsupervised free-living gait data in Parkinson's disease patients. *J Neuroeng Rehabil.* (2021) 18:93. doi: 10.1186/s12984-021-00883-7

46. Martindale CF, Christlein V, Klumpp P, Eskofier BM. Wearables-based multi-task gait and activity segmentation using recurrent neural networks. *Neurocomputing.* (2021) 432:250–61. doi: 10.1016/j.neucom.2020.08.079

47. Roth N, Küderle A, Prossel D, Gassner H, Eskofier BM, Kluge F. An inertial sensor-based gait analysis pipeline for the assessment of real-world stair ambulation parameters. *Sensors.* (2021) 21:6559. doi: 10.3390/s21196559

48. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. (2015) 521:436–44. doi: 10.1038/nature14539

49. Eskofier BM, Lee SI, Daneault J-F, Golabchi FN, Ferreira-Carvalho G, Vergara-Diaz G, et al. Recent machine learning advancements in sensor-based mobility analysis: deep learning for Parkinson's disease assessment. In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* (2016). p. 655–58. doi: 10.1109/EMBC.2016.7590787

50. Chiari L, Della Croce U, Leardini A, Cappozzo A. Human movement analysis using stereophotogrammetry: part 2: instrumental errors. *Gait Post.* (2005) 21:197–211. doi: 10.1016/j.gaitpost.2004.04.004

51. Miller A. Gait event detection using a multilayer neural network. *Gait Post.* (2009) 29:542–5. doi: 10.1016/j.gaitpost.2008.12.003

52. Kidziński Ł, Delp S, Schwartz M. Automatic real-time gait event detection in children using deep neural networks. *PLoS ONE.* (2019) 14:e0211466. doi: 10.1371/journal.pone.0211466

53. Filtjens B, Nieuwboer A, D'cruz N, Spildooren J, Slaets P, Vanrumste B. A data-driven approach for detecting gait events during turning in people with Parkinson's disease and freezing of gait. *Gait Post.* (2020) 80:130–6. doi: 10.1016/j.gaitpost.2020.05.026

54. Lempereur M, Rousseau F, Rémy-Néris O, Pons C, Houx L, Quellec G, et al. A new deep learning-based method for the detection of gait events in children with gait disorders: proof-of-concept and concurrent validity. *J Biomech.* (2020) 98:10941090. doi: 10.1016/j.jbiomech.2019.109490

55. Gadaleta M, Cisotto G, Rossi M, Rehman RZU, Rochester L, Del Din S. Deep learning techniques for improving digital gait segmentation. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. (2019). p. 1834–37. doi: 10.1109/EMBC.2019.8856685

56. Mazzà C, Alcock L, Aminian K, Becker C, Bertuletti S, Bonci T, et al. Technical validation of real-world monitoring of gait: a multicentric observational study. *BMJ Open.* (2021) 11:e050785. doi: 10.1136/bmjopen-2021-050785

57. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* (2005) 53:695–9. doi: 10.1111/j.1532-5415.2005.53221.x

58. Bertuletti S, Cereatti A, Comotti D, Caldara M, Croce UD. Static and dynamic accuracy of an innovative miniaturized wearable platform for short range distance measurements for human movement applications. *Sensors.* (2017) 17:1492. doi: 10.3390/s17071492

59. Bertuletti S, Croce UD, Cereatti C. A wearable solution for accurate step detection based on the direct measurement of the inter-foot distance. *J Biomech.* (2019) 84:274–7. doi: 10.1016/j.jbiomech.2018.12.039

60. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* (2011) 12:2825–30.

61. Salis F, Bertuletti S, Bonci T, Caruso M, Scott K, Alcock L, et al. A multi-sensor wearable system for the assessment of diseased gait in real-world conditions. *Front Bioeng Biotechnol.* (2023) 11:1143248. doi: 10.3389/fbioe.2023.1143248

62. Viceconti M, Penna SH, Dartee W, Mazzà C, Caulfield B, Becker C, et al. Toward a regulatory qualification of real-world mobility performance biomarkers in Parkinson's patients using digital mobility outcomes. *Sensors*. (2020) 20:5920. doi: 10.3390/s20205920

63. Salis F, Bertuletti S, Scott K, Caruso M, Bonci T, Buckley E, et al. A wearable multi-sensor system for real world gait analysis. In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. (2021). p. 7020–23. doi: 10.1109/EMBC46164.2021.9630392

64. Salis F, Bertuletti S, Bonci T, Della Croce U, Mazzà C, Cereatti A. A method for gait events detection based on low spatial resolution pressure insoles data. *J Biomech.* (2021) 127:11061187. doi: 10.1016/j.jbiomech.2021.110687

65. Rossanigo R, Caruso M, Salis F, Bertuletti S, Croce UD, Cereatti A. An optimal procedure for stride length estimation using foot-mounted magneto-inertial measurement units. In: *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. (2021). p. 1−6. doi: 10.1109/MeMeA52024.2021.9478604

66. Bertoli M, Cereatti A, Trojaniello D, Avanzino L, Pelosin E, Din SD, et al. Estimation of spatio-temporal parameters of gait from magneto-inertial measurement units: multicenter validation among Parkinson, mildly cognitively impaired and healthy older adults. *Biomed Eng Online.* (2018) 17:58. doi: 10.1186/s12938-018-0488-2

67. Bruijn SM, van Dieën JH, Meijer OG, Beek PJ. Statistical precision and sensitivity of measures of dynamic gait stability. *J Neurosci Methods.* (2009) 178:327–33. doi: 10.1016/j.jneumeth.2008.12.015

68. Rispens SM, van Schooten KS, Pijnappels M, Daffertshofer A, Beek PJ, van Dieën JH. Identification of fall risk predictors in daily life measurements: gait characteristics' reliability and association with self-reported fall history. *Neurorehabil Neural Repair.* (2015) 29:54–61. doi: 10.1177/1545968314532031

69. van Schooten KS, Rispens SM, Elders PJM, van Dieën JH, Pijnappels M. Toward ambulatory balance assessment: estimating variability and stability from short bouts of gait. *Gait Post.* (2014) 39:695–9. doi: 10.1016/j.gaitpost.2013.09.020

70. van Schooten KS, Pijnappels M, Rispens SM, Elders PJM, Lips P, van Dieën JH. Ambulatory fall-risk assessment: amount and quality of daily-life gait predict falls in older adults. *J Gerontol Ser A.* (2015) 70:608–15. doi: 10.1093/gerona/glu225

71. Warmerdam E, Hansen C, Romijnders R, Hobert MA, Welzel J, Maetzler W. Full-body mobility data to validate inertial measurement unit algorithms in healthy and neurological cohorts. *Data.* (2022) 7:136. doi: 10.3390/data7100136

72. Bai S, Zico Kolter J, Koltu V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271* (2018).

73. Rémy P. *Temporal Convolutional Networks for Keras*. (2021). Available online at: https://github.com/philipperemy/keras\discretionary-tcn

74. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (2016). p. 770−8. doi: 10.1109/CVPR.2016.90

75. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. In: *4th International Conference on Learning Representations (ICLR)* (2016).

76. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *Proceedings of the 32nd International Conference on Machine Learning*. PMLR (2015) p. 448–56.

77. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res.* (2014) 15:1929–58.

78. Bishop CM. *Pattern Recognition and Machine Learning*. Singapore: Springer Science+Business Media, LLC. (2006).

79. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. 1st ed. MIT Press (2016).

80. Pham MH, Elshehabi M, Haertner L, Del Din S, Srulijes K, Heger T, et al. Validation of a step detection algorithm during straight walking and turning in patients with Parkinson's disease and older adults using an inertial measurement unit at the lower back. *Front Neurol.* (2017) 8:457. doi: 10.3389/fneur.2017.00457

81. Bonci T, Salis F, Scott K, Alcock L, Becker C, Bertuletti S, et al. An algorithm for accurate marker-based gait event detection in healthy and pathological populations during complex motor tasks. *Front Bioeng Biotechnol.* (2022) 10:868928. doi: 10.3389/fbioe.2022.868928

82. Diez D, Çetinkaya-Rundel M, Barr C. *OpenIntro Statistics*. 4th ed. OpenIntro, Inc. (2019).

83. Del Din S, Godfrey A, Rochester L. Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and Parkinson's disease: toward clinical and at home use. *IEEE J Biomed Health Inform.* (2016) 20:838–47. doi: 10.1109/JBHI.2015.2419317

84. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *Statistician.* (1983) 32:307. doi: 10.2307/2987937

85. Giavarina D. Understanding bland altman analysis. *Biochem Med.* (2015) 25:141–51. doi: 10.11613/BM.2015.015

86. Ji N, Zhou H, Guo K, Williams Samuel O, Huang Z, Xu L, et al. Appropriate mother wavelets for continuous gait event detection based on time-frequency analysis for hemiplegic and healthy individuals. *Sensors.* (2019) 19:3462. doi: 10.3390/s19163462

87. Catalfamo P, Ghoussayni S, Ewins D. Gait event detection on level ground and incline walking using a rate gyroscope. *Sensors*. (2010) 10:5683–702. doi: 10.3390/s100605683

88. Del Din S, Elshehabi M, Galna B, Hobert MA, Warmerdam E, Suenkel U, et al. Gait analysis with wearables predicts conversion to Parkinson disease. *Ann Neurol.* (2019) 86:357–67. doi: 10.1002/ana.25548

89. König A, Klaming L, Pijl M, Demeurraux A, David R, Robert P. Objective measurement of gait parameters in healthy and cognitively impaired elderly using the dual-task paradigm. *Aging Clin Exp Res.* (2017) 29:1181–9. doi: 10.1007/s40520-016-0703-6

90. Flachenecker F, Gaßner H, Hannik J, Lee D-H, Flachenecker P, Winkler J, et al. Objective sensor-based gait measures reflect motor impairment in multiple sclerosis patients: reliability and clinical validation of a wearable sensor device. *Mult Scler Relat Disord.* (2020) 39:10191003. doi: 10.1016/j.msard.2019.101903

91. Mohan DM, Habib Khandoker A, Wasti SA, Alali S III, Jelinek HF, Khalaf K, et al. Assessment methods of post-stroke gait: a 690 scoping review of technology-driven approaches to gait characterization and analysis. *Front Neurol.* (2021) 12:650024. doi: 10.3389/fneur.2021.650024

92. Griškevičius J, Apanskiene V, ŽiŽiene J, Daunoravičiene K, Ovčinikova A, Kizlaitiene R, et al. Estimation of temporal gait parameters of multiple sclerosis patients in clinical setting using inertial sensors. In: *Proccedings of 11th International Conference Biomdlore*. Lithuania: VGTU Technika (2016). doi: 10.3846/biomdlore.2016.21

93. Gausden EB, Sin D, Levack AE, Wessel LE, Moloney G, Lane JM, et al. Gait analysis after intertrochanteric hip fracture: does shortening result in gait impairment? *J Orthopaed Trauma.* (2018) 32:554–58. doi: 10.1097/BOT.0000000000001283

94. Sivakumar A, Rickman M, Thewlis D. Gait biomechanics after proximal femoral nailing of intertrochanteric fractures. *J Orthopaed Res.* (2023) 41:862–874. doi: 10.1002/jor.25427

95. Rampp A, Barth J, Schülein S, Gaßmann KG, Klucken J, Eskofier BM. Inertial sensor-based stride parameter calculation from gait sequences in geriatric patients. *IEEE Trans Biomed Eng.* (2015) 62:1089–97. doi: 10.1109/TBME.2014.2368211

96. Karantonis DM, Narayanan MR, Mathie M, Lovell NH, Celler BG. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Trans Inform Technol Biomed.* (2006) 10:156–67. doi: 10.1109/TITB.2005.856864

97. Dijkstra B, Kamsma Y, Zijlstra W. Detection of gait and postures using a miniaturised triaxial accelerometer-based system: accuracy in community-dwelling older adults. *Age Ageing.* (2010) 39:259–62. doi: 10.1093/ageing/afp249

98. Iluz T, Gazit E, Herman T, Sprecher E, Brozgol M, Giladi N, et al. Automated detection of missteps during community ambulation in patients with Parkinson's disease: a new approach for quantifying fall risk in the community setting. *J Neuroeng Rehabil.* (2014) 11:48. doi: 10.1186/1743-0003-11-48

99. Ullrich M, Küderle A, Hannink J, Del Din S, Gaßner H, Marxreiter F, et al. Detection of gait from continuous inertial sensor data using harmonic frequencies. *IEEE J Biomed Health Inform.* (2020) 24:1869–78. doi: 10.1109/JBHI.2020.2975361

100. Micó-Amigo ME, Bonci T, Paraschiv-Ionescu A, Ullrich M, Kirk C, Soltani A, et al. Assessing real-world gait with digital technology? validation, insights and recommendations from the *Mobilise-D Consortium. J Neuroeng Rehabil.* (2023) 20:78. doi: 10.1186/s12984-023-01198-5

101. Lord SR, Bindels H, Ketheeswaran M, Brodie MA, Lawrence AD, Close JCT, et al. Freezing of gait in people with Parkinson's disease: nature, occurrence, and risk factors. *J Parkinsons Dis.* (2020) 10:631–40. doi: 10.3233/JPD-191813

102. Lindemann U, Zijlstra W, Aminian K, Chastin SFM, De Bruin ED, Helbostad JL, et al. Recommendations for standardizing validation procedures assessing physical activity of older persons by monitoring body postures and movements. *Sensors.* (2014) 14:1267–77. doi: 10.3390/s140101267