A Walk in the Dark

Inertial Measurement Unit-Based Gait Event Detection in Healthy and Neurological Cohorts

Robbin Romijnders

Kiel University
Christian-Albrechts-Universität zu Kiel

A Walk in the Dark

Inertial Measurement Unit-Based Gait Event Detection in Healthy and Neurological Cohorts

Dissertation zur Erlangung des akademischen Grades eines
Doktor der Ingenieurwissenschaften
(Dr.-Ing.)
der Technischen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Robbin Romijnders

aus

Heesch (Niederlande)

Jahr

2023

Kiel University
Christian-Albrechts-Universität zu Kiel

Prüfungskommission

Vorsitz:	Prof. Dr. Martina Gerken
Betreuung:	Prof. Dr.-Ing. Gerhard Schmidt
weitere Mitglieder:	Prof. Dr. med. Walter Maetzler (UKSH)
	Prof. Dr.-Ing. Jan Trieschmann

Datum der Einreichung:	30. März 2023
Datum der mündlichen Prüfung:	14. Juli 2023

Declaration

I hereby declare that:

a) the content and design of this thesis, entitled "*Inertial Measurement Unit-Based Gait Event Detection in Healthy and Neurological Cohorts*", is, – apart from my supervisor's guidance – all my own work and only using the sources that are listed;
b) this thesis, neither partially nor wholly, has been submitted as part of a doctoral examination procedure to another examining body, nor has it been published or submitted for publication;
c) this thesis has been prepared subject to the Rules of Good Scientific Practice of the German Research Foundation;
d) an academic degree has never been withdrawn from me.

Place, Date:

_____

Robbin Romijnders

Contributions to Prepared Manuscripts

The following manuscripts have already been published or have been submitted for publication:

- Romijnders, R., Warmerdam, E., Hansen, C., Welzel, J., Schmidt, G., & Maetzler, W. (2021). Validation of IMU-based gait event detection during curved walking and turning in older adults and Parkinson's Disease patients. *Journal of NeuroEngineering and Rehabilitation*, 18(1), 28. doi: 10.1186/s12984-021-00828-0
- **Proportional contributions:** *Conceptual design, 80%; Planning, 80%; Implementation, 100%; Preparation of the manuscript, 100%*

- Romijnders, R., Warmerdam, E., Hansen, C., Schmidt, G., & Maetzler, W. (2022). A Deep Learning Approach for Gait Event Detection from a Single Shank-Worn IMU: Validation in Healthy and Neurological Cohorts. *Sensors (Basel, Switzerland)*, 22(10), 3859. doi: 10.3390/s22103859
- **Proportional contributions:** *Conceptual design, 90%; Planning, 80%; Implementation, 100%; Preparation of the manuscript, 100%*

- Romijnders, R., Salis, F., Hansen, C., ..., Schmidt, G., Maetzler, W., & Mobilise-D consortium (2022). Ecological validity of a deep learning algorithm to detect gait events from real-life walking bouts in mobility-limiting diseases. *Submitted to Digital Health.*
- **Proportional contributions:** *Conceptual design, 90%; Planning, 80%; Implementation, 70%; Preparation of the manuscript, 100%*

Place, Date:

Place, Date:

Robbin Romijnders

Prof. Dr.-Ing. Gerhard Schmidt

Preface

I am greatly honored that I was offered the chance to pursue a doctoral degree. I am also really happy that it now seems it will finally come to an end. I have tried my best, but I also have to acknowledge others who have been of great help during the efforts of the past four years.

Zuerst möchte ich Walter danken für die Möglichkeit, Teil der Arbeitsgruppe Maetzler zu sein. Wir haben viele interessante Diskussionen, verstehen uns nicht immer, aber auf einer persönlichen Ebene gibt es vielleicht keinen besseren, empathischeren, Kollegen. Offiziell führt mein Promotionsweg über die Technischen Fakultät. Dafür danke ich Gerhard, der mich auch sehr herzlich in seine DSS-Team aufgenommen hat. Ich habe insgeheim mehr gelernt, als ich zugeben will, und ich schätze die offene Tür.

Natürlich bedanke ich mich auch bei meinen anderen Freunden und Kollegen vom UKSH und dem DSS-Team, z.B. Crispy, Corina, Petra, Kathrin und Kevin, die eher auf organisatorischer Ebene eine große Unterstützung sind. Zwei Personen möchte ich jedoch ganz besonders erwähnen, nämlich Elke und Clint. Elke, dank nogmaals dat ik altijd met mijn vragen omtrent het leven in Kiel, de inschrijving bij de uni of Keep Control een antwoord vond. Het heeft me enorm geholpen om snel te acclimatiseren in het zonnige noorden! Lieber Clint, für dich ziehe ich auch gerne die Socken hoch, und bedanke ich mich gerne bei dir für all die Tassen Kaffee und Grillabende. Es ist hilfreich, sich gemeinsam über die Entscheidungen der Chief Commander zu beschweren. Ich hoffe dass du auch etwas von mir gelernt hast und vielleicht mal gewinnen kannst mit Pfeile werfen.

Verder wil ik natuurlijk mijn familie bedanken voor alle steun en het vertrouwen dat ik dit tot een goed eind kan brengen! Pap en mam, dank voor het nog altijd ter beschikking stellen van een bed, als ik weer eens voorbijkom. Hoewel ik natuurlijk al een tijd uit huis ben, is er altijd nog een verschil tussen *thuis* en *thuis thuis*.

Al final, doy más que las gracias a Melissa. Nuestros caminos se cruzaron en Kiel y hasta hace poco nunca hubiera imaginado que llegaría a compartir amores y penas. Gracias por todos los momentos hermosos, por el sentimiento cálido, por escuchar cuando algo me preocupaba, pero sobre todo por simplemente estar ahí. El futuro es incierto, pero afrontémoslo juntos.

Robbin Romijnders
Kiel, July 2023

Summary

Walking impairments are common in elderly people and its prevalence increases with age. Walking impairments have devastating consequences and are associated with a loss of mobility, increased institutionalization, increased fall risk and decreased quality of life. Numerous disorders of both the central and peripheral nervous system can cause an impaired walking pattern. The objective quantification of walking is therefore of high clinical interest for clinicians, researchers and neurological patients.

Walking is made up from repetitive gait cycles, that can be divided in a stance phase, during which the foot is in contact with the ground, and a swing phase, during which the same foot is swinging forward. These phases are demarcated by gait events that are referred to as initial and final contact. The robust and accurate detection of these gait events is critical for any clinical gait analysis. Recent advances in wearable inertial sensor technology potentially allow the clinical gait analysis to shift to long-term continuous monitoring in the habitual environment. However, to date, the algorithms to extract gait events from inertial measurement unit (IMU) data have limited ecological validity as they have been validated mainly in clinical research settings with straight-line walking trials.

In this thesis a deep learning (DL)-based network is developed to determine gait events from IMU data from a shank- or foot-worn device. The DL network takes as input the raw IMU data predicts for each time step the probability that it corresponds to an initial or final contact. The algorithm is validated for walking at different self-selected speeds across multiple neurological diseases and both in clinical research settings and the habitual environment. The algorithms shows a high detection rate for initial and contacts, and a small time error when compared to reference events obtained with an optical motion capture system or pressure insoles.

Based on the excellent performance, it is concluded that the DL algorithm is well suited for continuous long-term monitoring of gait in the habitual environment.

Zusammenfassung

Gangstörungen sind bei älteren Menschen weit verbreitet, und ihre Prävalenz nimmt mit dem Alter zu. Gangstörungen haben weitgehende Folgen und gehen mit einem Verlust an Mobilität, einer zunehmenden Institutionalisierung, einem erhöhten Sturzrisiko und einer verminderten Lebensqualität einher. Zahlreiche Störungen sowohl des zentralen als auch des peripheren Nervensystems können zu einer Beeinträchtigung des Gehverhaltens führen. Die objektive Quantifizierung des Gehens ist daher von großem klinischen Interesse für Kliniker, Forscher und neurologische Patienten.

Das Gehen besteht aus sich wiederholenden Gangzyklen, die in eine Standphase, in der der Fuß den Boden berührt, und eine Schwungphase, in der derselbe Fuß vorwärts schwingt, unterteilt werden können. Diese Phasen werden durch Gangereignisse abgegrenzt, die als erster und letzter Fußkontakt bezeichnet werden. Die robuste und genaue Erkennung dieser Gangereignisse ist für jede klinische Ganganalyse von entscheidender Bedeutung. Jüngste Fortschritte in der Technologie der tragbaren Inertialsensoren ermöglichen es, die klinische Ganganalyse auf eine langfristige, kontinuierliche Überwachung in der hauslichen Umgebung umzustellen. Bislang haben die Algorithmen zur Extraktion von Gangereignissen aus Inertialmessdaten (*Inertial Measurement Unit*: IMU) jedoch nur eine begrenzte ökologische Gültigkeit, da sie hauptsächlich in klinischen Forschungsumgebungen mit geradlinigen Gehversuchen validiert wurden.

In dieser Arbeit wird ein *Deep Learning* (DL)-basiertes Netzwerk entwickelt, um Gangereignisse aus IMU-Daten von einem Gerät zu bestimmen, das am Schaft oder am Fuß getragen wird. Das DL-Netzwerk nimmt die IMU-Rohdaten als Eingabe und sagt für jeden Zeitschritt die Wahrscheinlichkeit voraus, dass es sich um einen ersten oder letzten Kontakt handelt. Der Algorithmus wurde für das Gehen bei verschiedenen selbstgewählten Geschwindigkeiten bei mehreren neurologischen Erkrankungen und sowohl in klinischen Forschungsumgebungen als auch in der gewohnten Umgebung validiert. Der Algorithmus zeigt eine hohe Erkennungsrate für erster und letzter Fußkontakt und einen geringen Zeitfehler im Vergleich zu Referenzereignissen, die mit einem optischen markerbasierte System oder Druckeinlagen ermittelt wurden.

Aufgrund der ausgezeichneten Leistung wird der Schluss gezogen, dass sich der DL-Algorithmus gut für die kontinuierliche Langzeitüberwachung des Gangs in der gewohnten Umgebung eignet.

List of Figures

List of Tables

Outline

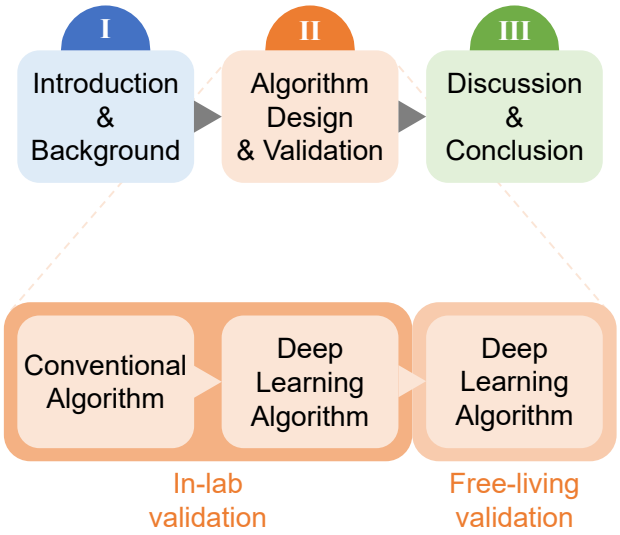


Figure 1: Schematic depiction of the thesis outline.

This thesis is divided into three main parts, and each part consists of one or more chapters (Figure 1).

Part I is comprised of the Introduction and Background chapters. In chapter 1 the topic of gait analysis and gait event detection is introduced. The importance of measuring gait is explained, which gait parameters are derived, and the aim of the current thesis is stated. In chapter 2, the signals that were used throughout this thesis are explained, and the signal processing methods are described.

Part II contains the contents from the works of the last years that have been published in, or submitted to, peer-reviewed journals. In chapter 3 a conventional gait event detection algorithm is validated for curved walking under single- and dual-task conditions. Then, in chapter 4 a deep learning-based algorithm is developed, and validated for its performance across different walking speeds and for different neurological and healthy cohorts. This deep learning algorithm is subsequently tested for its ecological validity across multiple mobility-limiting disease cohorts in real-life walking bouts (Chapter 5).

In the last part (Part III) the overall findings are summarized, the results are discussed, and suggestions are provided for further research.

Contents

Part I

Introduction & Background

1

Introduction

1.1. The Importance of Measuring Gait

Walking is the most common and functionally relevant aspect of mobility [Roc+20] and for short distances it is also the most convenient means of moving from one location to another [PB10]. Being able to walk forms a large part of functional independence, and therefore contributes greatly to social integration [HA05]. Furthermore, one's walking ability is closely related to participation (i.e., "the involvement in a life situation" [Wor02]), and thus to quality of life and well-being [Cuo+07; Car20]. The terms "walking" and "gait" are often used interchangeably, but strictly speaking gait refers to "the type or manner of walking" [LRW12]. In this work, the following definition is adopted [PB10; LRW12; Klu+21]:

"*Walking is a method of locomotion and is defined as initiating and maintaining a forward displacement of the centre of mass in an intended direction involving the use of the two legs, which provide both support and propulsion. The feet are repetitively and reciprocally lifted and set down whereby at least one foot is in contact with the ground at all times*".

As soon as parameters are extracted that tell about the walking pattern, the term "gait" is used. In that sense, "walking" refers more to the type of physical activity, whereas "gait" is associated with the assessment of that activity.

Walking requires a delicate balance between various interacting neuronal systems and virtually all levels of the nervous and musculoskeletal system are needed for walking [Nie03; Sni+07; Tak17] (Figure 1.1). Consequently, dysfunction in any of these systems can result in walking impairments [Sni+07]. Walking impairments commonly occur in elderly populations [Sni+07] and people with neurological disorders [War+20], and prevalence of these impairments increases with age

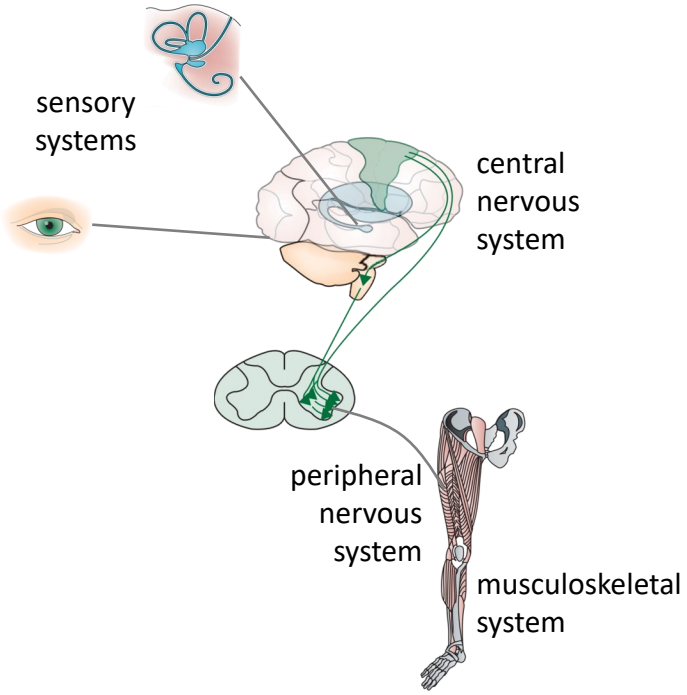


Figure 1.1: The systems involved in movement generation. *Adapted from [Sni+07] and based on [Car20].*

[PK17]. Markedly, about 20% of very old individuals (people aged 85 years or older) walk normally [Blo+92], suggesting that those who have walking impairments in fact suffer from an underlying disease [Sni+07].

Results from previous research indeed suggest that an altered walking pattern may reflect symptoms of an underlying cerebrovascular or neurodegenerative disease [Sni+07]. For example,

- in a population-based longitudinal study it was shown that the risk of cardiovascular death in subjects with senile gait disorders was twofold greater than in subjects with a normal gait [Blo+00],
- the time to walk 30 foot contributed independently to the prediction of time to onset of persistent cognitive impairment in a prospective, longitudinal, observational study [Mar+02],
- the presence of walking abnormalities in elderly persons without dementia at baseline is a significant predictor of the risk of development of dementia, especially non-Alzheimer's dementia [Ver+02],
- and quantitative gait parameters play an important role in identifying prodromal Parkinson's disease (PD) and progression within this phase [Del+19].

Taken together, walking is a sensitive indicator of overall health status which motivates the use of gait assessment for diagnostic and therapeutic purposes [PK17; Hod08].

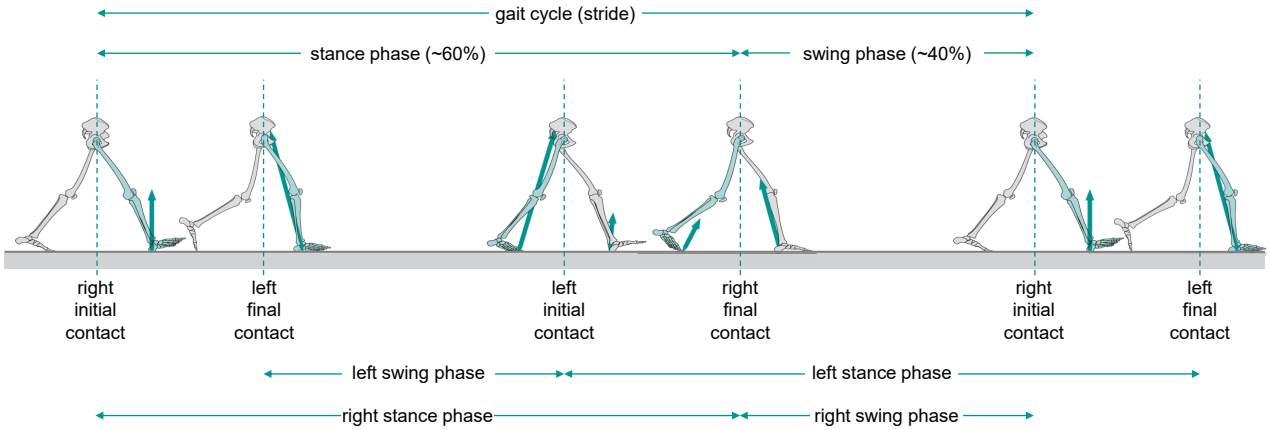


Figure 1.2: The gait cycle and associated gait phases and events. *Adapted from [LRW12].*

1.2. Gait Assessment

1.2.1. Gait Cycle

In order to understand walking impairments, it is important to first understand normal walking as this will provide the standard against which the impaired walking can be compared [LRW12; CS02].

Walking is a cyclic activity [Car20] and consists of repetitive gait cycles (Figure 1.2). A complete gait cycle is defined as the movement from the **initial contact (IC)** of one foot with the ground, to the next IC of the same foot [CS02; PB10; LRW12; SS20]. Mostly, the gait cycle is divided into two phases: stance and swing. The stance phase is the time during which the foot is in contact with the ground. The swing phase follows the stance phase and is the time during which the same foot is in the air. The separation of the two phases is marked by the **final contact (FC)** [PB10; LRW12; SS20]. A **step** is the time between an IC of the ipsilateral foot and the following IC of the contralateral foot [PB10; LRW12; Klu+21] and corresponds to a forward displacement of the foot together with a forward displacement of the trunk [DKZ10]. A **stride** is the time between two successive ICs of the same foot. As such, a stride is equivalent to the gait cycle and every stride contains a step from the ipsilateral foot and the consecutive step from the contralateral foot [PB10; LRW12; Klu+21].

There are many parameters that can be extracted from a gait analysis. Among the basic parameters are the spatio-temporal parameters such as step and stride times and lengths, and cadence (or step frequency) [LRW12; Ric18; Sol20]. Gait speed is also considered a key gait parameter [MFL15]. Subsequently derived parameters are usually considered secondary gait parameters, and serve as measures for gait symmetry [YB93; Sad+00; Ank+15], gait variability [Hau05], gait smoothness [LBM02; Hui+18], gait stability [Wol+85; Bru+09] and gait regularity [RM00; Ris+15; Hui+18].

1.2.2. Traditional Gait Assessment

Traditionally gait assessment is performed using tests or observations of the healthcare team in the clinical environment by applying, for example, the motor part of the Movement Disorder Society-sponsored Unified Parkinson's Disease Rating Scale (MDS-UPDRS, Part III; [Goe+08]) for patients with PD [Mae+21]. Observer-rated gait assessment is relatively subjective in nature and it has been suggested that this subjectivity may lead to poor validity, reliability, sensitivity, and specificity [TNF03], and depends on clinical expertise [Vie+10; Cel+21]. A more objective approach is to use timed walking tests [Pol+20], such as the six-minute walking test [But+82], short physical performance battery [Gur+94], 25-foot walking distance [Mot+17], and the timed up and go test [PR91]. However, both the rating scales and timed walking tests may not be sensitive to disease severity and cannot evaluate disease-specific characteristics [Muñ+19; BMM18; Sch+17; Mir+19]. Furthermore, these assessments are thought to mainly asses the gait capacity (i.e., How good is your [maximal] function?; [Mae+21]). This assessment approach does not tell anything about how well and in what quantity people are walking in daily life.

1.2.3. Instrumented Gait Assessment

Instrumented gait assessment using different digital-based technologies provides information that is complementary to clinical observation [Buc+19; Cel+21]. Different digital-based technologies are available for the objective instrumented assessment of gait, and the main consideration when deciding the best approach is the need to balance the requirement for better granularity, sensitivity, specificity, measurement accuracy, and minimal rater bias, with the complexity and feasibility of using such methods in clinics, communities, and clinical trials [Buc+19; Bon+20]. Three-dimensional (3D) motion capture (mocap) systems, instrumented walkway systems and force plates have been pioneering non-wearable systems that are considered the "gold standard" or reference for capturing kinetic or kinematic gait parameters with reasonable to high accuracy [Cel+21; Dic10; McD+01; MGM14]. Mocap systems use (passive) retro-reflective markers along with a video-based opto-electronic stereophotogrammetric system and various anatomical models to calculate the displacement of attached markers with high temporal and spatial accuracy [LRW12; Ric18; War+21; Cel+21; Chi+05]. However, the high costs, long preparation time, and need for specialist staff to operate the mocap systems form barriers to the adoption in routine clinical care [God+16; Buc+19]. Furthermore, there is increasingly more evidence that gait measured in laboratory settings under clinical supervision lacks ecological validity, that is, gait as measured in the lab does not reflect daily life gait [War+20; FM20; Hil+19; Atr+21].

1.2.4. Wearable Sensor-Based Gait Assessment

Recent years have therefore shown a clear trend towards wearable sensor-based gait analysis [War+20; Cel+21] and especially inertial measurement units (IMUs) have received much attention. IMUs allow for unobtrusive, long-term monitoring of gait in the habitual environment (e.g., the home, at work, or in the community) [Roc+20; Buc+19; War+20; FM20; Maz+21].

IMUs typically consist of 3D accelerometers, 3D gyroscopes and optionally 3D magnetometers, that measure linear accelerations, angular velocity and magnetic field strength, respectively. From these signals the timings of gait events, such as ICs and FCs, can be determined using dedicated algorithms [Pan+18; NK21]. These algorithms, however, are often only validated for straight-line walking in clinical environments, and were tested mostly in healthy populations. Gait event detection is then based on rule-based heuristics using empirically derived thresholds that do not necessarily translate to daily life gait, since IMU-based gait signals are affected by disease characteristics, participant activity levels and the exact context in which walking takes place [War+20; FM20; Pan+18; Pol+21]. However, recent advances in deep learning algorithms, computing resources and the availability of labeled gait data are paving the way for a data-driven approach for wearable sensor-based gait event detection [Lem+20; KDS19; Gad+19; Fil+20].

1.3. Aim of the Thesis

In summary, the objective quantification of the gait cycle is of high interest for different stakeholders, including patients, clinicians and researchers testing new treatment options, since gait is a surrogate marker of overall health status. The gait pattern may reflect the presence of an underlying neurodegenerative disease, and can inform on the future development of diseases like dementia and PD.

Using IMUs it is possible to record the gait pattern unobtrusively not just in specialized laboratories but also in the habitual environment. This provides a more complete picture of the mobility status and possible disease burden. However, the ecological validity of most algorithms is still an ongoing challenge, therefore the aim of the current thesis is **to develop and validate an algorithm for robust and accurate gait event detection based on IMU signals, which holds in both supervised and unsupervised settings across both healthy and different neurodegenerative disease cohorts**.

2

Background

2.1. Practical Considerations

IMUs are devices that typically contain a three-axis accelerometer and gyroscope [KHS17]. An accelerometer measures the external specific force acting on the device. The specific force consists of both the device's acceleration and the earth's gravity. A gyroscope measures the device's angular velocity, i.e., the rate of change of the device's orientation [KHS17]. Today, many gyroscopes and accelerometers are based on microelectromechanical system (MEMS) technology [KHS17]. For the purpose of measuring human motion, IMUs are strapped or attached to a body segment of interest. The recorded accelerations and angular velocities as measured by the device then represent the movement of the body segment it is attached to. Ideally, the sensor measurements can be used directly to inform on the body segment's pose, i.e., position and orientation, but inertial measurements are noisy and biased [KHS17; Sol20; Ios+16]. Thus, there a several practical considerations that need to be taken into account.

First, an IMU needs to be properly calibrated and drift should be accounted for [FGP95; Lam+16]. Furthermore, depending on the type of movement that is to be recorded, a convenient sampling frequency (f_s) and range is to be chosen [Sol20]. For walking it was shown that 98% of the power of the signal is contained below 10 Hz and 99% is below 15 Hz [AM85] and no significant acceleration frequency component was found above 16 Hz for either a low back-worn or a heel-worn IMU during treadmill walking [Ami+95]. Considering not just walking, but a range of daily life-relevant activities, it was found that the major energy band was 0.3 – 3.5 Hz [SH93]. Therefore, given that walking is our main activity of interest and to comply with the Nyquist-Shannon sampling theorem [SG14], the sampling frequency is typically 32 Hz or higher, and the range is typically around $\pm 8g$ for the

accelerometer and $\pm 2000^{\circ}$/s for the gyroscope to avoid saturation [All+19; Sol20].

2.2. Notation

The measurements from a single IMU that records acceleration and angular velocity at a given sampling frequency, f_{s}, are denoted:

$$\mathbf{x}(t_n) = \mathbf{x}[n] = \begin{bmatrix} a_{\mathsf{x}}[n] & a_{\mathsf{y}}[n] & a_{\mathsf{z}}[n] & \omega_{\mathsf{x}}[n] & \omega_{\mathsf{y}}[n] & \omega_{\mathsf{z}}[n] \end{bmatrix}^{\mathsf{T}} \tag{2.1}$$

for time $t_n = (n-1)/f_{\mathsf{s}}$ with $n = 1, \cdots, N$, and $a_i[n]$ and $\omega_i[n]$ denoting the accelerations and angular velocities, respectively, along the ith sensitive axis at the nth time step. The data for a complete measurement are denoted:

$$\mathbf{X} = \begin{bmatrix} — & (\mathbf{x}[1])^{\mathsf{T}} & — \\ — & (\mathbf{x}[2])^{\mathsf{T}} & — \\ & \vdots & \\ — & (\mathbf{x}[N])^{\mathsf{T}} & — \end{bmatrix} = \begin{bmatrix} a_{\mathsf{x}}[1] & a_{\mathsf{y}}[1] & a_{\mathsf{z}}[1] & \omega_{\mathsf{x}}[1] & \omega_{\mathsf{y}}[1] & \omega_{\mathsf{z}}[1] \\ a_{\mathsf{x}}[2] & a_{\mathsf{y}}[2] & a_{\mathsf{z}}[2] & \omega_{\mathsf{x}}[2] & \omega_{\mathsf{y}}[2] & \omega_{\mathsf{z}}[2] \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{\mathsf{x}}[N] & a_{\mathsf{y}}[N] & a_{\mathsf{z}}[N] & \omega_{\mathsf{x}}[N] & \omega_{\mathsf{y}}[N] & \omega_{\mathsf{z}}[N] \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_D \\ | & | & \cdots & | \end{bmatrix} \tag{2.2}$$

with $\mathbf{x}_d, d = 1, \cdots, D$ denoting the measurements for the dth sensor channel.

Similarly, if for a given measurement the gait events have been derived using a valid reference system, then the reference annotations (or labels) for the different classes (e.g., IC and FC) are denoted:

$$\mathbf{y}(t_n) = \mathbf{y}[n] = \begin{bmatrix} y_1[n] & \cdots & y_c[n] & \cdots & y_C[n] \end{bmatrix}^{\mathsf{T}} \tag{2.3}$$

where $c = 1, \cdots, C$, and with $C = 2$ in case only gait events from a single side are considered, and $C = 4$ in case gait events from both left and right side are considered simultaneously. For both cases it holds that $y_c[n] = 1$ if the data at time step n correspond to an event of the cth gait event class, and $y_c[n] = 0$ otherwise. Furthermore, an additional class label is introduced for those time steps that do not correspond to any of the gait events, and this class is referred to as the *null* class:

$$y_0[n] = \begin{cases} 1 & \text{if } y_c[n] = 0, \text{ for } c \in \{1, \cdots, C\} \\ 0 & \text{otherwise} \end{cases} \tag{2.4}$$

Combining Equations 2.3 and 2.4, the reference annotations at the nth discrete time step:

$$\mathbf{y}(t_n) = \mathbf{y}[n] = \begin{bmatrix} y_0[n] & y_1[n] & \cdots & y_c[n] & \cdots & y_C[n] \end{bmatrix}^{\mathsf{T}} \tag{2.5}$$

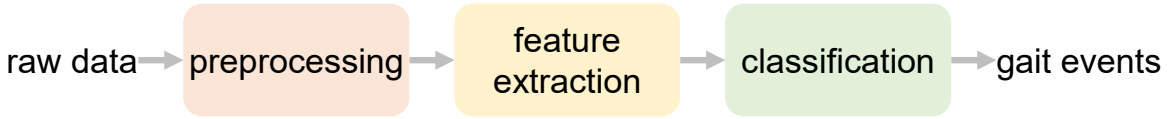


Figure 2.1: High-level overview of the signal processing steps to extract temporal gait parameters from raw IMU signals.

and for the complete measurement:

$$\mathbf{Y} = \begin{bmatrix} — & (\mathbf{y}[1])^\mathsf{T} & — \\ — & (\mathbf{y}[2])^\mathsf{T} & — \\ & \vdots & \\ — & (\mathbf{y}[N])^\mathsf{T} & — \end{bmatrix} = \begin{bmatrix} y_0[1] & y_1[1] & \cdots & y_C[1] \\ y_0[2] & y_1[2] & \cdots & y_C[2] \\ \vdots & \vdots & \vdots & \vdots \\ y_0[N] & y_1[N] & \cdots & y_C[N] \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_C \\ | & | & \cdots & | \end{bmatrix} \tag{2.6}$$

2.3. Conventional Approaches

Getting from the rather abstract measurements of accelerations and angular velocities to clinically relevant parameters requires accurate detection of the IC and FC [Rue+10]. Conventionally, a classical approach is adopted to idenfity ICs and FCs from the acceleration and angular velocity signals. This conventional approach consists of three main steps (Figure 2.1):

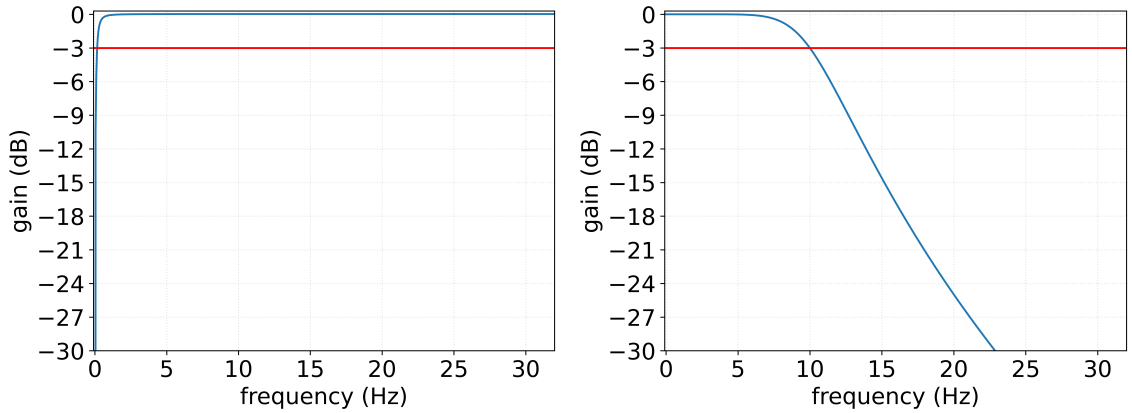
1. preprocessing the raw IMU signals,
2. extracting relevant features from the preprocessed signals,
3. and using the extracted features to classify IC or FC.

In this conventional approach, the feature extraction and classification are based on domain expert-defined rules and empirically determined thresholds.

2.3.1. Preprocessing

Preprocessing serves to reduce the effects of unwanted distortions that derive from low-frequency signal drift and/or high-frequency noise [BBS14; Fig+10; Sal+04; Mil+15; Del+16b]. In addition, IMU-based algorithms often require the sensors' orientation relative to an anatomical coordinate system defined according to the conventions of biomechanics (Figure 2.2; [Sol20; Car20]).

Figure 2.2: The signal processing steps that make up the preprocessing pipeline.



(a) High-pass filter.

(b) Low-pass filter

Figure 2.3: Filter characteristics of the applied high-pass and low-pass filters with the red horizontal line representing the -3 dB cut-off frequency.

Filtering

In order to reduce the effects of low-frequency signal drift one can subtract the channel-wise mean of the signal [Del+16b], subtract a linear fit of the signal (i.e., detrend the signal) [Pha+17] or apply a high-pass filter, for example using an infinite impulse response (IIR) filter [SG14; PV08]. For the presented studies, an IIR filter is used that is described by the following transfer function [Sal+04]:

$$H[z] = \frac{1 - z^{-1}}{1 - \alpha\, z^{-1}}, \quad \alpha = 0.995 \tag{2.7}$$

for a sampling frequency, f_s, of 200 Hz.

The high-pass filtered signals are then low-pass filtered where the most common method in IMU-based gait analysis is a 4th-order Butterworth filter with a cut off-frequency between 15 and 20 Hz. [Pha+17; Del+16b; Hic+16; God+08]. Both these filters (Figure 2.3) are applied twice to the signal, once forwards and once backwards, to achieve zero phase delay [KJ74; Sal+04].

Alignment

Gait assessment is based on the principles of biomechanics and by convention movement is described related to an anatomical coordinate frame [WC95; Wu+02]. Normally three anatomical planes are defined from the coordinate system (Figure 2.4; [Car20; Sol20]):

- a sagittal plane containing the upward and forward axes,
- the frontal or coronal plane containing the upward and lateral axes,
- and the transverse plane containing the forward and the lateral axes.

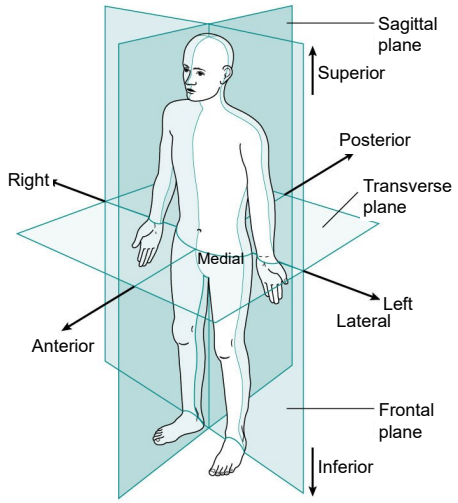


Figure 2.4: The anatomical planes and axes. (Adopted from [LRW12].)

As the IMU is attached to the participant's body segment by means of elastic straps or adhesive tape, the device orientation or alignment is normally offset with respect to the anatomical axes [Mil+15; Del+16b; Sol20; Car20]. The procedure to determine the orientation difference between the sensor and the body segment is referred to as sensor-to-body, or sensor-to-segment, alignment, and is crucial to derive any clinically meaningful angular kinematics data [Pac+20; LGA19]. The majority of studies perform a functional alignment, where a certain pose or known movement is performed such that at least one anatomical axis can be estimated [Pac+20; Fav+09; Kon+16a; de +10]. Other studies assume that the sensor is pre-aligned with an axis of interest, make use of an underlying kinematic model, or use a source other than the IMU [Pac+20; LGA19].

For the present studies it is assumed that at least one of the axes aligns with the anatomical medio-lateral axis, and no sensor-to-segment calibration procedure or signal processing technique is used.

2.3.2. Feature Extraction

Gait Event Detection

For the current studies, a conventional algorithm is implemented that detects gait events from the medio-lateral angular velocity signal [Sal+04]. It is based on first identifying local maxima in the medio-lateral angular velocity signal that correlate to midswing of the swing phase of the gait cycle [Ami+02], both for healthy [Sab+05] and pathological gait [Sal+04; Tro+14]. Local minima before and

after the midswing-associated maximum are then considered FC and IC, respectively (Figure 2.5; [Ami+02; Sal+04; Tro+14]).

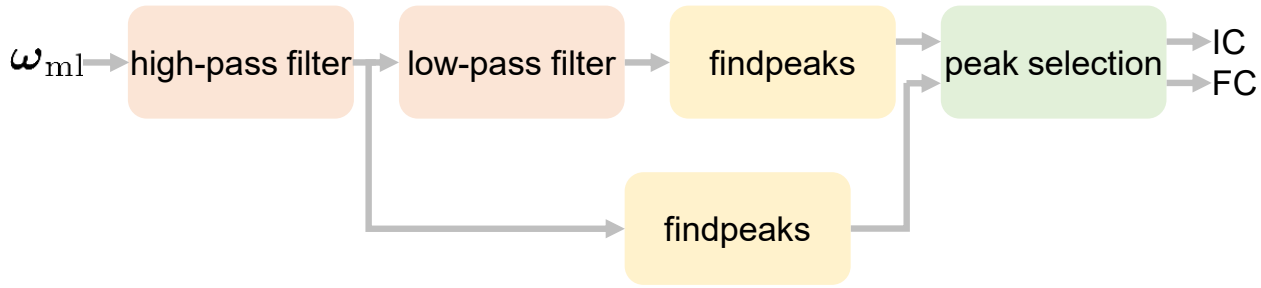


Figure 2.5: Signal flowchart for detecting gait events, i.e. initial contact (IC) and final contact (FC) from the raw medio-lateral angular velocity, ω_{ml}, based on local minima surrounding the peaks that correspond to midswing (MS).

For the current studies, the medio-lateral angular velocity, $\omega_{\mathrm{ml}}(t)$, is first high-pass filtered using the filter from Equation 2.7, and then low-pass filtered using a 4th-order Butterworth filter with the -3 dB cut-off frequency at 10 Hz. Next, local maxima and minima are detected using SciPy [Vir+20]. Timings of midswing are identified as local maxima with a minimum value of 50°/s in the band-pass filtered signal [Sal+04; Car+19]. If multiple adjacent local maxima within a distance of 500 ms are detected, the peak with the highest amplitude is selected and the rest is discarded [Sal+04]. Likewise, local minima with a value < -20°/s are detected in the high-pass filtered signal [Sal+04]. Finally, the last local minimum before each midswing and the first local mininum after each midswing are identified as the FC and IC, respectively (Figure 2.6; [Ami+02; Sal+04; Tro+14; Car+19]).

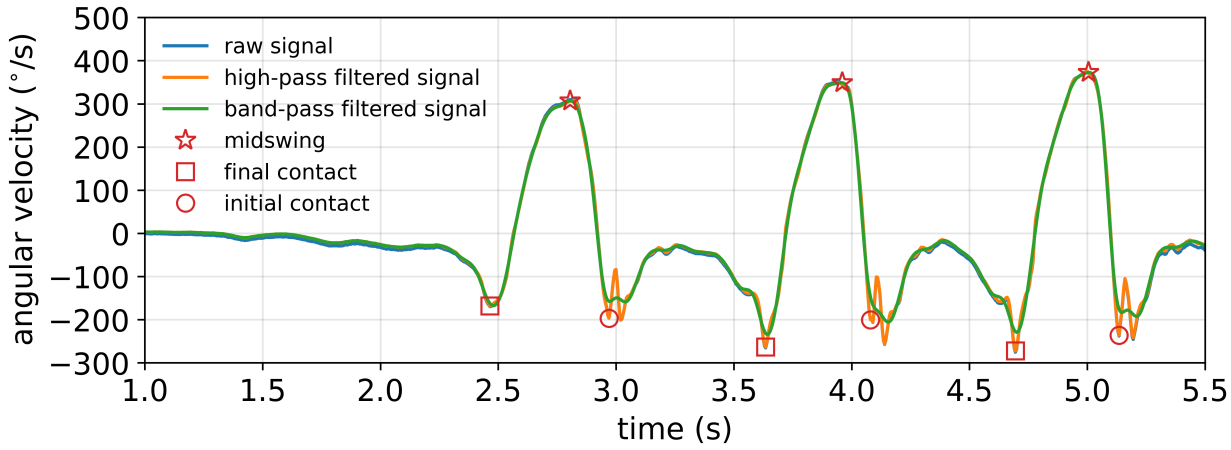


Figure 2.6: Detection of initial and final contacts from the medio-lateral angular velocity signal. The data are from a 5 meter walking trial of a healthy younger adult at self-selected preferred walking speed.

2.4. Deep Learning Approaches

A major drawback of this conventional approach is that it relies on handcrafted features (e.g., local maxima, local minima) and the classification depends on empirically determined thresholds (e.g., minimum peak height, time interval between consecutive peaks). The gait pattern, as recorded with an IMU, depends on the body segment to which it is attached, but also on contextual factors, such as age, mood, personality, and sociocultural factors (does one live in a rather urban or rural area) [PK17]. This implies that gait event detection performance might change for walking in different contexts [War21]. Furthermore, the feature extraction assumes that the sensor signals align with pre-defined anatomical axis, and this cannot always be guaranteed for unsupervised monitoring settings where study participants may have to attach IMUs themselves, for example after night rest or showering.

Alternatively, a deep learning (DL)-based approach can be adopted. The idea behind DL is that the model or network learns valuable representations from the data itself, and no explicit rules are defined for feature extraction or classification [Van18; Mei20; Cho21]. This approach is potentially invariant to orientation misalignment, and is transferable and scalable to other body segments. Furthermore, it is a suitable approach for a variety of time series-associated segmentation and classification tasks [Lön+19].

2.4.1. Convolutional Neural Networks

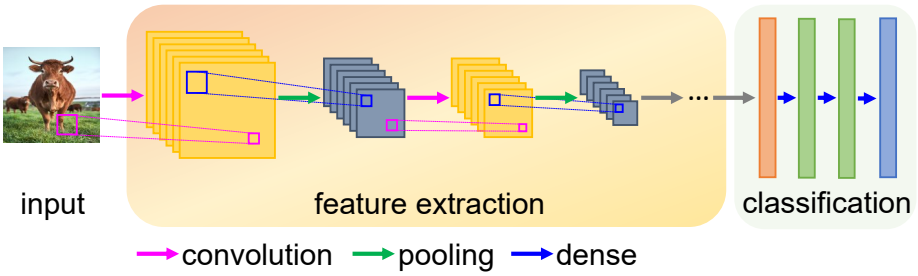
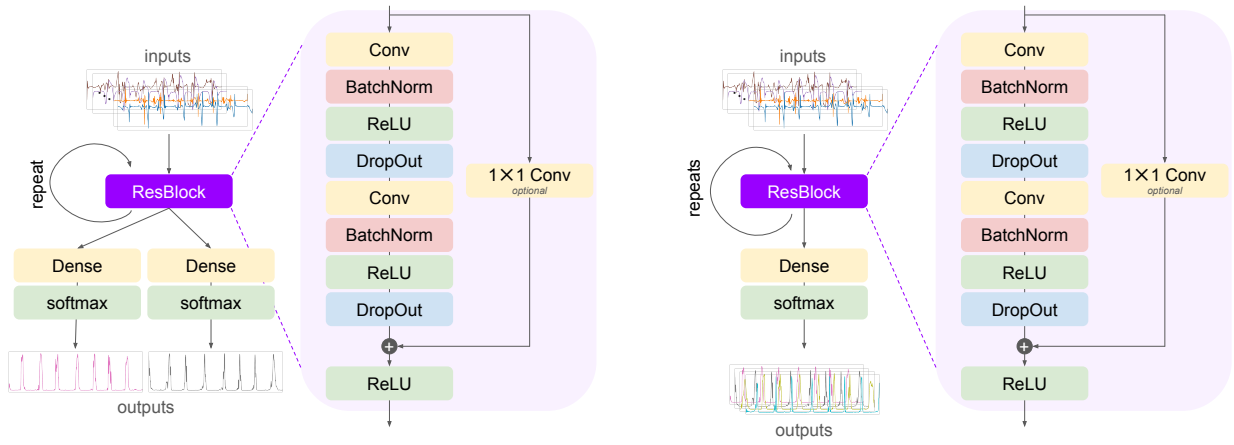


Figure 2.7: A typical convolutional neural network architecture. It consists of stacked layers that perform convolutions, pooling (or subsampling), flattening and non-linear activation. Based on [MS22].

To exploit the sequential nature of the gait pattern, a logical choice for a DL algorithm is either a recurrent neural network (RNN) or a convolutional neural network (CNN). Recent studies have found that CNNs based on dilated convolutions outperform RNNs for most sequence modelling tasks [BKK18; Li+22]. These specific CNNs are sometimes referred to as temporal convolutional networks (TCNs) [BKK18; van+16a; YK16; YKF17].

CNNs were first reported about in the early 1980s [Fuk80] and evolved over the course of time into what is now referred to as a CNN [LBH95]. Later, convolutional and pooling layers were introduced



(a) Network architecture for Chapter 4 with data from a single side considered, and the gait events modelled as two separate outputs.

(b) Network architecture for Chapter 5 with data from both left and right side considered simultaneously.

Figure 2.8: Temporal convolutional network (TCN) models for gait events detection with a separate and independent output for each gait event class (left) and with a single multi-class output (right).

[LeC+98], that together with fully-connected (or dense) layers and non-linear activation functions form the building blocks of nearly all CNN architectures [Cho21] (Figure 2.7).

2.4.2. Temporal Convolutional Networks

For the current thesis, a TCN is developed and validated for gait event detection from IMU data (Chapters 4 and 5). The TCN consists of stacked residual blocks, that are made up from two sequences of a (1) convolutional layer, (2) batch normalization layer, (3) rectified linear unit (ReLU) layer, and (4) an optional dropout layer, after which the output is added to a residual connection (either an identity mapping or a 1×1 convolutional layer) and a final ReLU activation layer is applied (Figure 2.8).

The mathematical processing of the input data, $\mathbf{X}$, through the consecutive layers of the residual blocks is described in detail in Appendix A. How the outputs of the last residual block are treated depends on whether the data from both sides are considered separately (thus allowing for a single sensor setup, like in Chapter 4) or considered simultaneously (like in Chapter 5).

- *Chapter 4*. In this case, the inputs, $\mathbf{X}$, consist of the raw acceleration and angular velocity data of a single IMU from either the left or right side. The labels are the ICs and FCs of the corresponding side, and these are considered two separate outputs (Figure 2.8a). The outputs of the last residual block are therefore passed through two separate dense (fully-connected) layers, and subsequently through a softmax layer to give the predicted events. The problem is cast as a regression problem, and the difference between the predictions and the labels are quantified by a weighted mean squared error (MSE) [KDS19; Fil+20]. The overall error is then

the average of the sum over the time steps, classes and (training) examples.

- *Chapter 5.* In this case, the inputs, **X**, consist of the raw acceleration and angular velocity data of the IMUs on both the left and right side. The labels are the left and right ICs and the left and right FCs. The outputs of the last residual block are passed through a single dense layer and subsequently through a softmax layer to give the predicted events. The problem is cast as a multi-class classification problem, and the difference between the predictions and the labels are quantified by a weighted categorical cross-entropy.

Part II

Algorithm Design & Validation

3

Gait Event Detection During Curved Walking and Dual-Task Conditions

The contents of this chapter are based on:

Romijnders, R., Warmerdam, E., Hansen, C., Welzel, J., Schmidt, G., & Maetzler, W. (2021). Validation of IMU-based gait event detection during curved walking and turning in older adults and Parkinson's Disease patients. *Journal of NeuroEngineering and Rehabilitation*, 18(1), 28. doi: 10.1186/s12984-021-00828-0

Abstract

Background

Identification of individual gait events is essential for clinical gait analysis, because it can be used for diagnostic purposes or tracking disease progression in neurological diseases such as Parkinson's disease. Previous research has shown that gait events can be detected from a shank-mounted inertial measurement unit (IMU), however detection performance was often evaluated only from straight-line walking. For use in daily life, the detection performance needs to be evaluated in curved walking and turning as well as in single-task and dual-task conditions.

Methods

Participants (older adults, people with Parkinson's disease, or people who had suffered from a stroke) performed three different walking trials: (1) straight-line walking, (2) slalom walking, (3) Stroop-and-walk trial. An optical motion capture system was used a reference system. Markers were attached to the heel and toe regions of the shoe, and participants wore IMUs on the lateral sides of both shanks. The angular velocity of the shank IMUs was used to detect instances of initial foot contact (IC) and final foot contact (FC), which were compared to reference values obtained from the marker trajectories.

Results

The detection method showed high recall, precision and F1 scores in different populations for both initial contacts and final contacts during straight-line walking (IC: recall = 100%, precision = 100%, F1 score = 100%; FC: recall = 100%, precision = 100%, F1 score = 100%), slalom walking (IC: recall = 100%, precision ≥ 99%, F1 score =100%; FC: recall = 100%, precision ≥ 99%, F1 score =100%), and turning (IC: recall ≥ 85%, precision ≥ 95%, F1 score ≥91%; FC: recall ≥ 84%, precision ≥ 95%, F1 score ≥89%).

Conclusions

Shank-mounted IMUs can be used to detect gait events during straight-line walking, slalom walking and turning. However, more false events were observed during turning and more events were missed during turning. For use in daily life we recommend identifying turning before extracting temporal gait parameters from identified gait events.

3.1. Introduction

Gait is recognized as a surrogate marker of health, and provides essential clinical insights in neurological disease status [DCM18; Buc+19]. Traditionally, gait has been assessed by visual observation, which suffers from subjectivity and imprecision [TNF03]. To overcome these limitations, multi-camera optical motion capture (OMC) systems can be used, but these systems are relatively expensive and restricted to expertise laboratories [Ios+16]. Furthermore, there is increasing evidence that the gait pattern observed in clinical gait assessments does not reflect daily-life gait [Hil+19; War+20]. Hence, to get a more complete picture of health status, there is an increasing demand for methods that allow for long-term gait monitoring in ambulatory settings. Inertial measurement units (IMUs) provide a promising alternative to assess gait in an objective, unobtrusive and unconstrained manner [Ios+16; Haj+18].

The term “gait” refers to “the way of walking” [Rue+10; LRW12] and human gait is commonly segmented into repetitive gait cycles. A normal gait cycle begins and ends with initial contact (IC), the instance when the foot strikes the ground [PB10]. The time interval between two consecutive ICs of the same foot is referred to as the gait cycle time or stride time. The time interval between two successive ICs of the opposite feet is called the step time. If, additionally, the event of final foot contact (FC) is considered, then all phases in the gait cycle can be described: swing and stance phase, or single and double support phase [DCM18; PB10]. Identification of gait events (GEs) and phases is considered essential for clinical gait assessment [Rue+10]. GEs can be detected from a single low back-mounted IMU [Par+19; Pha+17; McC+12; ZH03; Zij04; TCD14], however findings suggest that detecting GEs is easier from shank- or foot-mounted IMUs [Sal+04; Tro+14; Kon+16b] where foot-mounted IMUs increase errors, especially in pathological gait patterns [Kon+16b; LYL11].

The performance of IMU-based GE detection is, however, often tested only with treadmill walking [Pha+17; ZH03] or from walking trials where only the straight-line segments of walking trajectories were included in the analysis [Zij04; Sal+04; DGR16]. For more complex walking tasks, such as slalom walking or dual-task walking, one often relies on visually counting of the number of steps, which does not allow to assess the time error of the GE detection and is more prone to errors. Whether IMU-based GE detection is still valid in more complex walking tasks is yet to be shown. Daily-life gait is likely influenced by obstacle negotiation (approximately $30\,\%$ of daily-life gait is spent along curved trajectories [Gla+07; Tur+18]) and dual-/multi-tasking [Hil+19].

The aim of this study is therefore to quantify the performance of IC and FC detection in straight-line walking under single-task and dual-task conditions, and to quantify detection performance in curved walking and turning in (healthy) older adults (OA), people diagnosed with Parkinson’s disease (PD), and people who have suffered from a stroke (ST).

3.2. Methods

A step was considered as the interval between consecutive ICs of the ipsi- and contralateral foot [PB10], and corresponding to forward displacement of the foot together with a forward displacement of the trunk [DKZ10]. A stride was the interval between two consecutive ICs of the same foot, and as such it was equivalent to the gait cycle and every stride consisted of two steps [Rue+10; PB10].

3.2.1. Study Population

Table 3.1: Demographics data of the study participants summarized by group. N is the number of participants (between brackets the number of female participants). Age, height, mass, UPDRS-III and disease duration are presented as mean (standard deviation). Disease duration is the time since first diagnosis for PD, and time since stroke for ST. OA: older adults, PD: Parkinson's disease, ST: stroke, UPDRS-III: Unified Parkinson's Disease Rating Scale, part III: Motor Examination.

Group	N	**Age**	**Height**	**Mass**	**UPDRS-III**	**Disease duration**
		years	m	kg		years
OA	11 (2)	71 (9)	1.76 (0.07)	78.5 (13.5)	4 (3)	
PD	14 (5)	64 (10)	1.78 (0.08)	91.3 (14.7)	29 (21)	9 (5)
ST	9 (2)	68 (10)	1.75 (0.08)	81.3 (18.0)	6 (9)	2 (4)

Three different groups were distinguished: (1) OAs with no signs of any movement disorders, (2) PD participants in the medication ON state, and (3) ST participants (Table 3.1). For the OAs the minimum age was 60 years. All participants needed to be able to walk independently with or without walking aids. Exclusion criteria were a high fall risk (i.e. $>$ 2 falls in the last month, as reported by the participant), any impairment that refrained the participant from giving consent to participate in the study, and a score below 20 for the Montreal Cognitive Assessment (MoCA) [Nas+05]. All participants gave written informed consent and the study was approved by the ethical committee of the medical faculty at University Hospital Schleswig-Holstein (UKSH), No: D438/18.

3.2.2. Study Protocol

Participants walked a 5-meter distance that was marked at the start and end with two cones, approximately 1 meter apart (Figure 3.1). Participants were asked to start walking approximately two steps before the start, and to stop walking approximately two steps after the end. For the analysis of GE detection, we only considered the events that were registered within the 5-meter distance.

The following walking trials were performed:

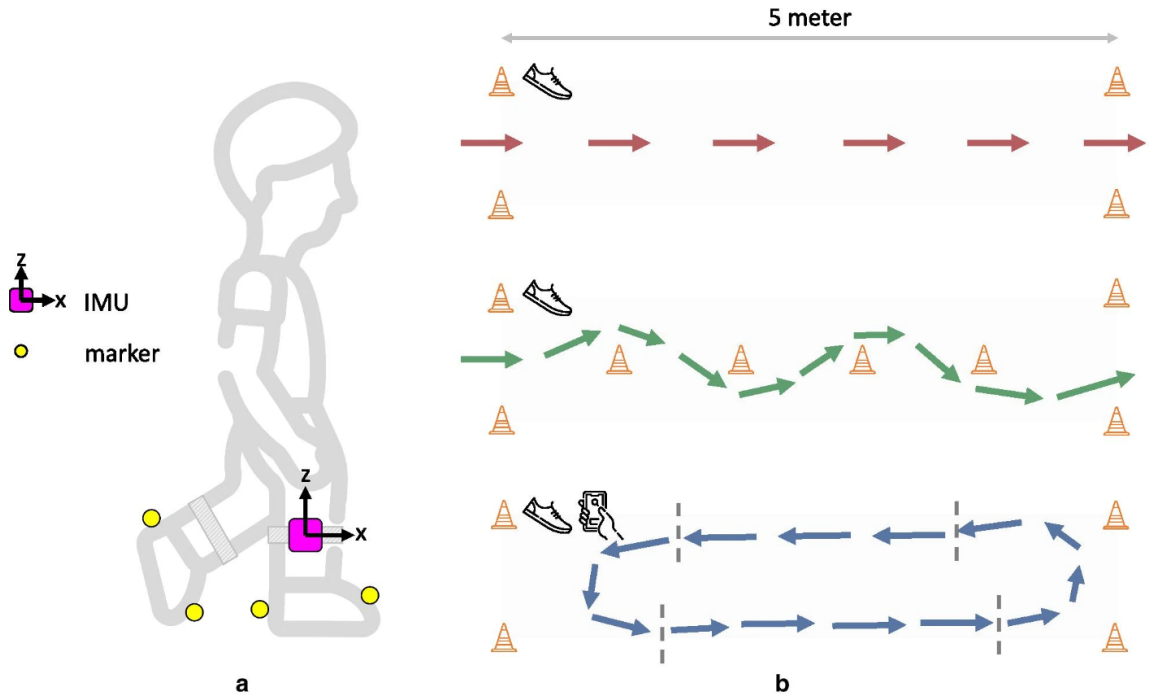


Figure 3.1: Schematic illustration of the setup. **a** Participants were equipped with inertial measurement units (IMUs) attached to the lateral sides of the shanks and reflective markers that were attached to the heel and toe region of the shoes. **b** Three different walking trials were performed: (top) straight-line walking trial, (middle) slalom walking trial, and (bottom) Stroop dual-task walking trial. Straight-line walking and slalom walking were performed only under single-task conditions (indicated by the shoe icon), whereas the Stroop task was as a cognitive-motor dual-task (indicated by the shoe and mobile phone icons). For the Stroop-and-walk trial, we distinguish between steps during straight-line segments (within the dashed vertical lines) and steps during turns (outside of the dashed vertical lines).

- straight-line trial, 5 meter, at preferred speed,
- slalom trial, 5 meter with a cone at every 1 meter, at preferred speed,
- Stroop-and-walk trial, walking up and down the 5-meter distance, while performing a numerical Stroop test [HT82] on a hand-held mobile phone until completion of the Stroop test, at preferred speed. For the numerical Stroop test, two numbers were displayed on the mobile screen that were different in value and different in semantic size. The participant needed to tap the number with the highest value. No further instructions as to prioritize any task were given.

For the Stroop-and-walk trial, participants started within the 5-meter distance, and walked up and down whilst turning on either end of the 5-meter distance (see Fig. 3.1 for an example trajectory). Turns were annotated manually using the Qualisys Track Manager 2018.1 software (QTM; Qualisys AB, Göteborg, Sweden), and GEs during turns were analyzed separately.

3.2.3. Optical Motion Capture System

Equipment

Reflective markers (diameter: 19 mm) were attached to the heel and toe of both left and right shoes (Fig. 3.1). Marker trajectories were recorded by a 12-camera optical motion capture system (Qualysis AB, Göteborg, Sweden) sampling at 200 Hz.

Signal Processing

GEs were detected from the heel marker trajectories, and provided reference values for the GE timings of ICs and FCs. Event timings were based on specific signal features from the heel marker vertical velocity and acceleration, respectively. Raw marker data were loaded into MATLAB (MATLAB 2018b, The Mathworks, Natwick, USA). The raw marker data were first interpolated to fill any gaps [GF16] and subsequently low-pass filtered with a 4th order Butterworth filter with a cut-off frequency, f_{cut}, of 5 Hz. The filter was applied to the marker data by using MATLAB's built-in filtfilt function, such that the filtered signal was not delayed.

The timings of ICs correlated with timings of local maxima in the vertical acceleration [HM00], whereas the timings of FCs closely correlated with timings of local maxima in the heel marker vertical velocity [PBv01]. Like in [Car+18], GEs were checked manually QTM.

3.2.4. Inertial Measurement Units

Equipment

One IMU (Noraxon U.S.A. Inc., Scottsdale, Arizona, USA) was attached to each shank with elastic straps. The sampling frequency was also set to 200 Hz and the OMC and IMUs were synchronized using a trigger at the beginning of each measurement [Qua11].

Signal Processing

Sensor data were loaded into MATLAB. IMUs were aligned such that the sensitive axes pointed roughly in antero-posterior direction (forward being positive), medio-lateral direction (left being positive), and the vertical direction (up being positive). Angular velocity was high-pass filtered using an IIR filter with $f_{\text{cut}} \approx 0.15$ Hz to reduce the effect of drift [Sal+04], and then low-pass filtered using a 4th order Butterworth filter with $f_{\text{cut}} = 10$ Hz. Both filters were applied using the `filtfilt` function.

Detection of GEs using shank-mounted IMUs was based on identifying negative peaks in the medio-lateral angular velocity that was high- and low-pass filtered. These negative peaks closely correlated to timings of mid-swing [Sal+04; Tro+14; Ami+02]. Only negative peaks with a value ≤ 10 % of the global minimum angular velocity were considered. Furthermore, if two or more consecutive peaks were detected within a time interval of 300 ms of each other, then only the peak with the lowest value was preserved [DGR16; Naj+03].

Data from the two legs were analyzed independently of one another to facilitate a setup with an IMU on a single side.

3.2.5. Data Analysis

Two GEs, IC and FC, were extracted using the reference system as well as the shank-mounted IMUs. For both events the detection performance was evaluated in terms of correctly identified events (true positives, TP), falsely identified events (false positives, FP) and missed events (false negatives, FN). TPs were defined as < 300 ms difference (in terms of magnitude) between an event detected by the IMU-based algorithm and the reference event [Pha+17]. From these metrics the recall, precision and F1 score were derived:

$$\text{recall} = R = \frac{\text{TP}}{\text{TP} + \text{FN}}$$
$$\text{precision} = P = \frac{\text{TP}}{\text{TP} + \text{FP}}$$
$$\text{F1 score} = F1 = 2\frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Recall expressed how many of the gait events were detected and precision expressed how many of the detected gait events were true gait events. The F1 score can be considered as a weighted average of the recall and precision. Furthermore, algorithm performance was evaluated by assessing the time error between the reference event (from the marker-based algorithm) and the predicted event [KW16; Ji+19], defined as:

$$\epsilon_{\text{IC}} = t_{\text{IC}} - t'_{\text{IC}}$$
$$\epsilon_{\text{FC}} = t_{\text{FC}} - t'_{\text{FC}}$$

where t_{IC} and t_{FC} denoted the time of the predicted IC and FC from the IMU-based algorithm, and t'_{IC} and t'_{FC} denoted the reference time of the IC and FC obtained from the OMC.

The effect of dual-task conditions was investigated by comparing GE detection from the straight-line trial to the GE detection from the straight-line segment of the Stroop-and-walk trial (Fig. 1). The effect of curved walking was investigated by comparing GE detection from the straight-line walking to slalom walking and turns from the Stroop-and-walk trial.

3.2.6. Statistical Analysis

Detection of Gait Events

The algorithm performance in detection of GEs was evaluated by generating contingency tables and comparing the recall, precision and F1 scores.

Time Agreement

Time agreement was assessed by determining the mean time error, corresponding $95\,\%$ confidence intervals and the mean absolute error (MAE). Confidence intervals (CIs) were computed as $\bar{x} \pm 1.96s$ with $\bar{x}$ the mean time error, and s the standard deviation of the time errors.

Comparison of time errors between tasks for each group. A Wilcoxon signed-rank test [Wil45] was used to compare each subject's mean values of the absolute errors for the single-task versus the dual-task conditions, and similarly for the straight-line walking, slalom walking, and turns [Tro+14]. Differences were considered statistically significant if the p-value was less than 0.05.

Comparison of time errors between groups for each task. A Wilcoxon rank sum test was used to compare the subject mean values of the absolute errors from the OA group and those obtained for the PD and ST group [Tro+14]. Differences were considered statistically significant if the p-value was less than 0.05.

3.3. Results

3.3.1. Detection of Gait Events

Effect of Dual-Tasking

Table 3.2: Validation results of gait event detection for the straight-line segments of the single-task trial and the Stroop-and-walk trial. F1: F1 score, FN: false negative, FP: false positive, OA: older adults, P: precision, PD: Parkinson's disease, R: recall, ST: stroke, TP: true positive.

		Initial contacts						**Final contacts**					
	N	**TP**	**FN**	**FP**	**R** %	**P** %	**F1** %	**TP**	**FN**	**FP**	**R** %	**P** %	**F1** %
Straight-line trial													
OA	11	83	0	0	100	100	100	83	0	0	100	100	100
PD	14	131	0	0	100	100	100	133	0	0	100	100	100
ST	9	78	0	0	100	100	100	81	0	0	100	100	100
Stroop-and-walk trial (straight-line segments													
OA	11	501	5	0	99	100	100	497	8	1	98	100	99
PD	11	587	1	1	100	100	100	589	2	2	100	100	100
ST	9	457	12	0	97	100	99	451	18	2	96	100	98

IMU-based GE detection showed high recall (IC: $\geq$97 %, FC: $\geq$96 %), high precision (IC: $\geq$100 %, FC: $\geq$100 %) and high F1 score (IC: $\geq$99 %, FC: $\geq$98 %) for the three different groups in both single-task and dual-task conditions for GEs during straight walking (Table 3.2). All ICs and FCs were detected for the single-task trials. In the straight walking segments from the Stroop-and-walk trials a number of false events (1 IC, 5 FC) and a number of missed events (18 IC, 28 FC) were observed.

Effect of Curved Walking and Turns

Table 3.3: Validation results of gait event detection during straight-line walking, curved walking, and turns. F1: F1 score, FN: false negative, FP: false positive, OA: older adults, P: precision, PD: Parkinson's disease, R: recall, ST: stroke, TP: true positive.

		Initial contacts						**Final contacts**					
	N	**TP**	**FN**	**FP**	**R** %	**P** %	**F1** %	**TP**	**FN**	**FP**	**R** %	**P** %	**F1** %
Straight-line trial													
OA	11	83	0	0	100	100	100	83	0	0	100	100	100
PD	14	131	0	0	100	100	100	133	0	0	100	100	100
ST	9	78	0	0	100	100	100	81	0	0	100	100	100

Table 3.3: Validation results of gait event detection during straight-line walking, curved walking, and turns. F1: F1 score, FN: false negative, FP: false positive, OA: older adults, P: precision, PD: Parkinson's disease, R: recall, ST: stroke, TP: true positive.

		Initial contacts						**Final contacts**					
	N	**TP**	**FN**	**FP**	**R** %	**P** %	**F1** %	**TP**	**FN**	**FP**	**R** %	**P** %	**F1** %
Slalom trial													
OA	11	103	0	0	100	100	100	96	0	0	100	100	100
PD	14	181	0	1	100	99	100	190	0	1	100	99	100
ST	9	124	0	0	100	100	100	126	0	0	100	100	100
Stroop-and-walk trial (turns)													
OA	11	297	51	11	85	97	91	296	55	15	84	95	89
PD	11	312	45	18	87	95	91	315	42	14	88	96	92
ST	9	281	41	15	87	95	91	282	39	13	88	96	92

IMU-based GE detection showed high recall (IC: $100\,$%, FC: $100\,$%), high precision (IC: $\geq 99\,$%, FC: $\geq 99\,$%) and high F1 score (IC: $99\,$%, FC: $98\,$%) for the three different groups for the straight-line walking and slalom walking (Table 3.3). All ICs and FCs were detected for the single-task trials. In the straight walking segments from the Stroop-and-walk trials a number of false events (1 IC, 5 FC) and a number of missed events (18 IC, 28 FC) were observed. One IC and one FC were falsely detected for a single PD patient in the slalom walking trial, where the patient swung its foot forward and backward without taking a step (that is, there was a swing phase but the patient did not move forward but rather put its foot down on the same spot). For ICs and FCs during turns, for the three groups, recall was lower than for the straight-line and slalom walking (IC: $\geq 85\,$%, FC: $\geq 84\,$%), and likewise for the precision (IC: $\geq 95\,$%, FC: $\geq 95\,$%) and F1 score (IC: $\geq 91\,$%, FC: $\geq 89\,$%). More events were missed by the IMU-based gait event detection (137 ICs, 136 FCs) and more false events were detected (44 ICs, 42 FCs).

3.3.2. Time Agreement

Effect of Walking Task on Gait Event Detection

Effect of Dual-Task Walking.

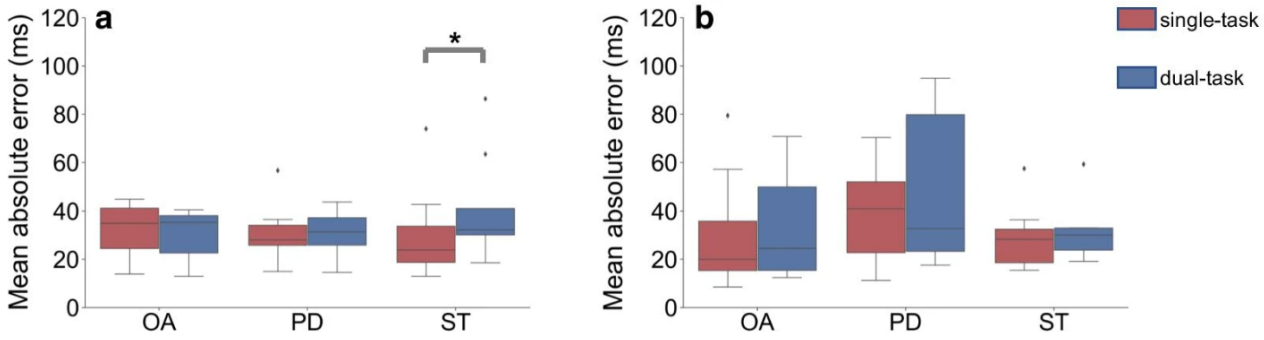


Figure 3.2: Boxplots showing the mean absolute errors for the **a** initial contacts and **b** final contacts detection for the gait events during straight-line walking under single-task (red) or dual-task (blue) conditions. *: $p < 0.05$

Table 3.4: Values for the time errors of the gait events. CI: confidence interval, MAE: mean absolute error, OA: older adults, PD: Parkinson's disease, sd: standard deviation, ST: stroke.

		Initial contacts			**Final contacts**		
		mean (**sd**)	**MAE**	$95\,\%$**CI**	**mean** (**sd**)	**MAE**	$95\,\%$**CI**
		ms	ms	[ms, ms]	ms	ms	[ms, ms]
OA	straight-line	14 (36)	32	[−57, 85]	−25 (33)	31	[−90, 39]
	stroop-and-walk	21 (33)	32	[−44, 86]	−21 (38)	33	[−96, 54]
PD	straight-line	26 (29)	33	[−30, 83]	−25 (45)	40	[−113, 63]
	stroop-and-walk	11 (40)	33	[−67, 89]	−43 (49)	51	[−139, 52]
ST	straight-line	17 (36)	31	[−54, 87]	−6 (33)	29	[−71, 59]
	stroop-and-walk	31 (43)	41	[−53, 116]	−4 (38)	32	[−79, 71]

Table 3.4 shows the mean time errors, mean absolute error (MAE) and the $95\,\%$ CI for the GEs during straight-line walking, either from the straight-line trial (single-task) or the Stroop-and-walk trial (dual-task). A Wilcoxon signed-rank test showed that there are no significant differences between the single-task and dual-task conditions, except for the ST group (Figure 3.2). In the ST group we found that the MAE is significantly higher in dual-task conditions compared to single-task conditions (p $= 0.039$, $W = 5.0$).

Effect of Curved Walking and Turns.

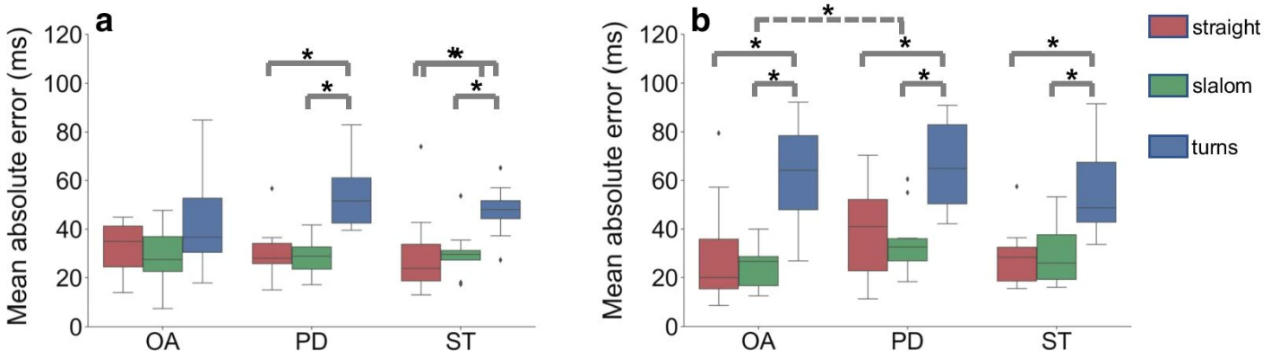


Figure 3.3: Boxplots showing the mean absolute errors for the **a** initial contacts and **b** final contacts detection for the gait events during straight-line walking (red), slalom walking (green) and during turns (blue). *: $p < 0.05$

Table 3.5: Values for the time errors of the gait events. CI: confidence interval, MAE: mean absolute error, OA: older adults, PD: Parkinson's disease, sd: standard deviation, ST: stroke.

		Initial contacts			**Final contacts**		
		mean (sd)	**MAE**	**95 %CI**	**mean (sd)**	**MAE**	**95 %CI**
		ms	ms	[ms, ms]	ms	ms	[ms, ms]
OA	straight-line	14 (36)	32	[−57, 85]	−25 (33)	31	[−90, 39]
	slalom	18 (32)	30	[−45, 81]	−2 (31)	25	[−63, 59]
	stroop-and-walk (turns)	11 (57)	40	[−101, 123]	−33 (77)	58	[−183, 117]
PD	straight-line	26 (29)	33	[−30, 83]	−25 (45)	40	[−113, 63]
	slalom	20 (33)	30	[−44, 83]	−15 (45)	35	[−103, 72]
	stroop-and-walk (turns)	−11 (81)	56	[−170, 148]	−45 (75)	68	[−192, 103]
ST	straight-line	17 (36)	31	[−54, 87]	−6 (33)	29	[−71, 59]
	slalom	19 (36)	31	[−53, 90]	−4 (34)	28	[−72, 63]
	stroop-and-walk (turns)	12 (62)	47	[−110, 133]	−7 (77)	54	[−158, 144]

Table 3.5 shows the mean time errors, MAE and 95 % CI for the GEs during straight-line walking, slalom walking and during turns. A Wilcoxon signed-rank test showed that for all groups there are no significant differences between the straight-line walking and the slalom walking, for both ICs and FCs. It is also shown that in IC detection significant differences exist for the MAE of the straigt-line

walking and turns, and the slalom walking and turn, for both the PD and ST group (Figure 3.3). For the FC detection, significant differences were observed for the MAEs between straight-line walking and turns, and slalom walking and turns, for all groups.

Effect of Group on Gait Event Detection

A Wilcoxon rank-sum test showed that no significant differences were observed between groups for all walking tasks, except for slalom walking, where in FC detection the MAE is significantly larger in the PD group compared to the OA group ($p = 0.039$, $W = -2.608$).

3.4. Discussion

In this study, shank-mounted IMU-based detection of gait events was tested for different walking tasks and in different mobility-limiting chronic diseases. The detection performance was evaluated for GEs from steps during straight-line walking under both single-task and dual-task conditions. Furthermore, the detection performance was evaluated for GEs from steps during straight-line walking, curved walking and turns. Three different groups of participants were distinguished (OA, PD, ST; Table 3.1) to evaluate whether the detection performance is affected by presence or history of neurological disease.

The high (i.e., almost perfect) recall, precision and F1 score (Table 3.2) imply that IMU-based detection of ICs and FCs is feasible in both single-task and dual-task straight-line walking for participants with physiological and different pathological walking patterns. Similarly, the high (i.e., again almost perfect) recall, precision and F1 score (Table 3.3) imply that IMU-based gait event detection is feasible in curved walking by assessing the detection performance in slalom walking. Again, results hold for participants across different walking conditions. The performance of detecting gait events during turning shows lower recall, precision and F1 score which suggests that shank-mounted IMUs are less feasible to detect GEs during turning. It should be noted that this holds only for how currently the signals were processed. If the vertical acceleration signal would also have been used, like in [Maq+16] for lower limb amputees, then GE detection would likely have been less dependent on the (forward) swinging motion of the leg.

Concerning the differences between IMU-based event timings and reference event timings, results were in a range similar to previous studies [Car+18; Ji+19]. There are many possible contributors to the time difference, and possibly a combination of these will be in play. First, and most importantly, two different systems were used and distinct signal characteristics were used to identify the same event. For the reference system, local extrema in the marker velocity and acceleration marked the instances of the event [HM00; PBv01]. For the IMUs, local extrema in the angular velocity about the medio-lateral axis marked the instances of gait events [Sal+04; Tro+14; Ami+02]. More recent research found that there was no clear feature from the angular velocity signal related

to FC, at least when walking on a treadmill [Böt+16]. Next, the filtering of the angular velocity signal may contribute to the time error, as with a lower cut-off frequency less signal details were preserved which affects the presence of local extrema [CGE10]. Potentially, there is also a minor contribution to the time error from the hardware-triggered synchronization [Qua11], as mentioned before [Car+18; Sta+90; Chi+05]. Most importantly, for all groups and walking tasks, the $95\,\%$ CI of the estimated event timing encompassed $0\,$s and therefore the current methods are considered valid for detection of gait events [Sal+04; Tro+14; Ami+02]. The $95\,\%$ CI is largest for the time errors of IC and FC during turns.

It was found that the absolute time error is not significantly different when comparing detection of GEs for single-task conditions to dual-task conditions, except for the detection of ICs for the ST group. This may be explained by the altered gait pattern that can be observed in some post-stroke patients, especially while executing a cognitive-motor dual-task [Bae+13; Tim+18]. The mean absolute time error was found significantly higher for IC and FC detection during turns compared to both straight-line and slalom walking for the PD and ST group. Together with the lower recall, precision, F1 score and the larger $95\,\%$ CI this implies that a shank-mounted IMU is less feasible for detecting GEs during turning. What are the most probable reasons for this observation? In straight-line walking, the leg is swung forward reaching a peak angular velocity at approximately mid-swing [Sal+04; Tro+14; Ami+02]. The local maxima right before and after the peak are then correlated to the instances of FC and IC [Sal+04; Tro+14; Ami+02]. However, for steps during turning, this swinging motion may not always be observed, depending on which turning strategy is used [HS99] which may explain the higher number of FN and FP events during turns, compared to straight-line and, especially, slalom walking. The study has in our opinion a particular clinical relevance, considering that increasingly home-derived data will be used for patient management [War+20] and assessment in clinical trials [Cor+19; Cer18]. IMUs are ideally suited for this performance-based assessment. Our study suggests that the temporal components necessary for the qualitative assessment of gait (IC and FC) can be detected very reliably during straight and slalom walking (e.g. go for a stroll or shopping, commuting to work), but gait phases with rotations of e.g. 180° and possibly interrupted forward movement can be detected less reliably. This implies that turns should definitely be included in unsupervised IMU-based gait detection. Of course, this statement only applies to the sensor constellation as used in this study, and not, for example, to data from IMUs positioned at the low back. Although methods derived from low back- or foot-mounted IMUs [Pha+17; McC+12; LYL11; Ram+15] may be less susceptible to turns, literature suggested that GE detection is easier and more robust from shank-mounted IMUs [Sal+04; Tro+14; Kon+16b].

Furthermore, to continue improving the current methods, and to be less dependent on the forward swinging motion of the leg, future research may also include vertical acceleration signals [McC+12; Maq+16] or include information from both the time and frequency domain [KW16; Cai+17].

One of the limitations of this study is the relatively short walking distance. However, the focus of our research was on detecting GEs regardless of the walking distance. Another limitation is that the results are from a supervised assessment in a controlled environment, which is not representative of daily-life conditions [Hil+19; War+20].

3.5. Conclusion

Shank-mounted IMUs can be used to detect gait events from steps during straight-line and curved walking, under both single-task and dual-task conditions, in different neurological populations. Gait events from steps during turns can be detected but result in more missed events and more false events. In case spatio-temporal parameters are subsequently derived, the higher number of missed and false events will have a negative effect on these parameters. If turns are not automatically identified, the spatio-temporal parameters from ambulatory assessment should be interpreted with care.

Acknowledgements

The authors would like to thank Arash Atrsaei for providing feedback on the submission details.

Funding

Open Access funding enabled and organized by Projekt DEAL. This research was funded by Keep Control from the EU's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 721577.

Contributions

Robbin Romijnders (RR) performed the study, developed the algorithms, and analyzed the data. Elke Warmerdam (EW), Clint Hansen (CH) and Walter Maetzler (WM) designed the study protocol. EW conducted data collections. RR and Gerhard Schmidt (GUS) discussed the algorithms. WM found the financial support. RR, EW, CH, Julius Welzel (JW), GUS and WM discussed the results and revised the manuscript. All authors read and approved the final manuscript.

Ethics approval

The study was approved by the ethical committee of the medical faculty at the University Hospital Schleswig-Holstein (UKSH), No: D438/18.

4

Deep Learning for IMU-Based Gait Events Detection

The contents of this chapter are based on:

Romijnders, R., Warmerdam, E., Hansen, C., Schmidt, G., & Maetzler, W. (2022). A Deep Learning Approach for Gait Event Detection from a Single Shank-Worn IMU: Validation in Healthy and Neurological Cohorts. *Sensors (Basel, Switzerland)*, 22(10), 3859. doi: 10.3390/s22103859

Abstract

Background

Many algorithms use 3D accelerometer and/or gyroscope data from inertial measurement unit (IMU) sensors to detect gait events (i.e., initial and final foot contact). However, these algorithms often require knowledge about sensor orientation and use empirically derived thresholds. As alignment cannot always be controlled for in ambulatory assessments, methods are needed that require little knowledge on sensor location and orientation, e.g., a convolutional neural network-based deep learning model.

Methods

Therefore, 157 participants from healthy and neurologically diseased cohorts walked 5 m distances at slow, preferred, and fast walking speed, while data were collected from IMUs on the left and right ankle and shank. Gait events were detected and stride parameters were extracted using a deep learning model and an optoelectronic motion capture (OMC) system for reference. The deep learning model consisted of convolutional layers using dilated convolutions, followed by two independent fully connected layers to predict whether a time step corresponded to the event of initial contact (IC) or final contact (FC), respectively.

Results

Results showed a high detection rate for both initial and final contacts across sensor locations (recall ≥92%, precision ≥97%). Time agreement was excellent as witnessed from the median time error (0.005 s) and corresponding inter-quartile range (0.020 s). The extracted stride-specific parameters were in good agreement with parameters derived from the OMC system (maximum mean difference 0.003 s and corresponding maximum limits of agreement (−0.049 s, 0.051 s) for a 95% confidence level).

Conclusions

Thus, the deep learning approach was considered a valid approach for detecting gait events and extracting stride-specific parameters with little knowledge on exact IMU location and orientation in conditions with and without walking pathologies due to neurological diseases.

4.1. Introduction

Gait deficits are common in older adults and possibly reflect the presence of an underlying neurodegenerative disease [Sni+07; Hod08]. For example, conversion to Parkinson's Disease [Del+19] or conversion from mild cognitive impairment to Alzheimer's Disease [Kön+17; Ber+18] are linked with changes in spatiotemporal gait parameters. Similarly, temporal gait parameters are different for stroke patients [von+95; Moh+21] and patients with multiple sclerosis [Gri+16; Fla+19] when compared to healthy controls. To objectively quantify gait deficits, stride-specific parameters such as stride time or stride length are often used [Han+16]. The beginning and end of a stride are determined from two successive initial contacts (ICs) of the same foot [PB10; LRW12]. The IC is when the foot contacts the ground and together with the instant at which the foot leaves the ground (final contact, FC), each stride can be divided in a stance and swing phase [Rue+10; BR14]. The events of IC and FC, also referred to as *gait events*, are commonly determined using force or pressure measuring devices [BR14], or marker-based optoelectronic motion capture systems (OMC; henceforth referred to as the *marker-based* system or method) [Chi+05; TR20]. These systems are relatively expensive, and restricted to usage in expertise laboratories [Ios+16; Jar+18]. As there is increasing evidence that gait measured in the lab does not reflect daily-life gait [Hil+19; War+20; Atr+21], there is increasingly more interest in measurement systems that allow for continuous gait analysis in ambulatory settings. Therefore, the use of inertial measurement units (IMUs) is especially attractive, as these can be used to measure gait in ecologically valid environments, such as the home environment, thereby painting a more complete picture of health status [Del+16b; Sha+20] and providing clinical information that is complementary to standardized lab-based assessments [War+20; FM20; Atr+21; Cor+21].

Previous research suggests that gait event detection is more accurate using an IMU worn on a lower limb (e.g., shank or foot) compared to an IMU worn on the low back [BRG15; SBM16; Pan+18]. Now, in order to get from abstract IMU sensor readings to clinically relevant gait parameters (i.e., from accelerations and angular velocities to stride times) [Han+16], different algorithmic approaches have been developed in the last twenty years of clinical gait research. A recent study has evaluated a cross-section of these algorithms for different sensor locations on the lower leg and foot [NK21]. The algorithms were categorized according to which signals were analyzed, for example the angular velocity about the medio-lateral axis, or the accelerations along vertical and antero-posterior axes. This means the sensor readings need to be linked with the anatomical axes, that is, one needs to know which sensor axis aligns with for example the medio-lateral axis. In most approaches, it is simply assumed that due to sensor attachment the sensor axis aligns roughly with the anatomical axis of interest [Sal+04; Sab+05; Jas+06; CGE10; Tro+14; Maq+16; Rom+21] or an additional calibration procedure (e.g., [FGP95]) is required [Gre+10; NK21]. In ambulatory assessments however, study

participants often attach the sensor themselves, for example after showering, and therefore the sensor location and alignment cannot be controlled for. Furthermore, it is unlikely that each time the sensor is (re-)attached study participants, especially those with gait deficits, perform a calibration procedure that usually consists of holding a pre-defined pose and performing some known movement sequences [LGA19; Pac+20].

Taken together, this drives the need for an approach that is invariant to sensor orientation, and is applicable across a variety of pathological gait patterns. In the field of image analysis similar requirements have been successfully addressed by algorithms that share a common underlying methodology referred to as deep learning [LBH95; Han+16; ter19]. Recent applications of deep learning algorithms have already shown improved performance in detecting gait events from marker-based motion capture when compared to conventional, often heuristics-based, algorithms [KDS19; Lem+20; Fil+20]. Another study used a deep learning approach to detect gait events from either three IMUs (worn on the low back, and both ankles) or a single IMU (worn on the low back), and showed that the time error was considerably smaller for the deep learning algorithm than for a commonly applied wavelet-based approach [Gad+19].

The aim of this study is to extend these works by validating this deep learning approach for detecting gait events in a heterogeneous cohort of healthy and neurologically diseased adults considering a single IMU setup, worn on the lower leg.

4.2. Methods

4.2.1. Data Collection

Gait analyses were performed in the Universitätsklinikum Schleswig-Holstein (UKSH) campus Kiel, Germany. The study [War+21] was approved by the ethical committee of the medical faculty at the UKSH (no: D438/18). In total, data from 160 participants were included for the current analysis, including data from young adults (YA; age: 18 - 60 years), older adults (OA; age: $>$60 years), people with Parkinson's Disease (PD; according to the UK Brain Bank criteria [GL88], people with a recent (<4 weeks) symptomatic stroke (stroke), people with multiple sclerosis (MS; according to the McDonalds criteria [Tho+18]), people with chronic low back pain (cLBP), and people with diagnoses not fitting in any aforementioned groups or disorders with no explicit diagnosis (other) (Table 4.1). Inclusion criteria were an age of 18 years or older and the ability to walk independently without a walking aid. Participants were excluded from the study with a Montreal Cognitive Assessment [Nas+05] score $<$15 and other movement disorders that affected mobility, as noticed by the clinical assessor.

Table 4.1: Demographics data of the study participants summarized by group and gender. Age, height, and weight are presented as mean (standard deviation).

Group	Gender	Num. of participants	Age	Height	Mass
			years	cm	kg
YA	F	21	27 (7)	173 (5)	67 (9)
	M	21	29 (9)	185 (8)	80 (12)
OA	F	12	70 (6)	167 (6)	72 (17)
	M	10	73 (6)	180 (6)	83 (12)
PD	F	12	67 (6)	168 (7)	70 (15)
	M	19	61 (11)	178 (7)	86 (14)
MS	F	12	37 (10)	174 (9)	75 (9)
	M	9	42 (16)	189 (9)	96 (32)
stroke	F	4	66 (11)	160 (7)	65 (13)
	M	17	67 (18)	178 (7)	84 (15)
cLBP	F	3	67 (6)	168 (7)	65 (6)
	M	6	66 (17)	177 (8)	86 (14)
other	F	3	60 (16)	166 (4)	79 (19)
	M	8	68 (19)	182 (7)	85 (14)

Participants performed three walking trials consisting of walking 5 m at either (1) preferred speed ("*Please walk at your normal walking speed.*"), (2) slow speed ("*Please walk at half of your normal walking speed.*"), or (3) fast speed ("*Please walk as fast as possible, without running or falling.*"). The 5 m distance was marked with two cones on both ends, and participants were asked to start walking approximately two steps before the cones on one end, and stop walking approximately two steps after passing the cones on the other end.

For the current analysis, data from four IMUs (Noraxon USA Inc., myoMOTION, Scottsdale, AZ, USA) were considered, namely those that were attached laterally above the left and right ankle joint and those attached proximally at the left and right shank. IMUs were secured to participants using elastic bands with a special hold for the IMU. Furthermore, reflective markers were attached on top of the usual foot wear at the heel and toe of both feet (Figure 4.1). Marker data were recorded using a twelve-camera OMC system (Qualisys AB, Göteborg, Sweden) at a sampling frequency of 200

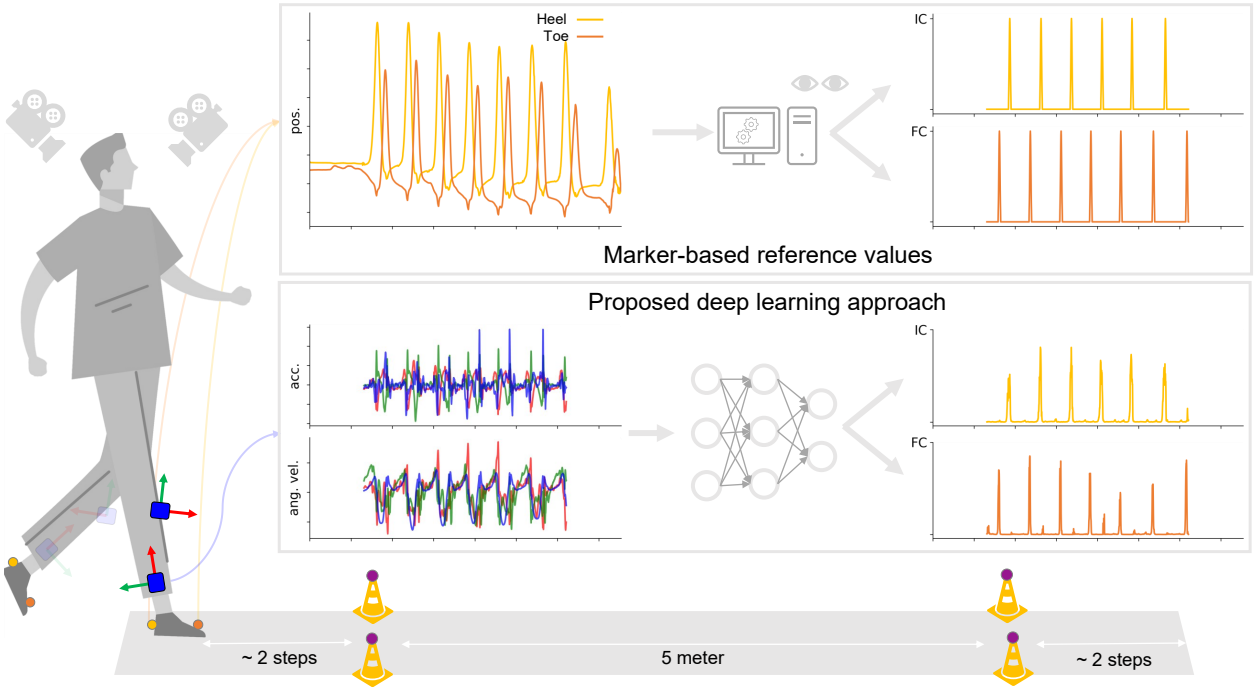


Figure 4.1: Schematic depiction (from `https://www.vecteezy.com/free-vector/man-walking`, accessed on 11 Nov 2021) of the current study. Study participants wore IMUs on the ankle and shanks, and reflective markers were adhered on the heel and toe of usual footwear (illustrated on the **left**). Marker data were used to obtain reference values for the timings of initial and final contacts (**top**), where accelerometer and gyroscope data from each tracked point were inputted to a neural network that predicted timings of the same initial and final contacts (**bottom**).

Hz. IMU data were recorded at the same sampling frequency, and both systems were synchronized using a TTL signal [War+21].

4.2.2. Data Preprocessing

Marker Data

First, data from both marker and IMU systems were cropped so that only data from within the 5 m distance were considered. Any gaps in the marker data were filled by interpolation making use of inter-correlations between markers [Fed13; GF16]. The data were then low-pass filtered using a sixth-order Butterworth filter with a cut-off frequency of 20 Hz [RK21]. The filter was applied twice to the input data [KJ74]. After filtering in the forward direction, the filtered sequence was reversed and run back through the filter [Sal+04]. The filtered data were differentiated to get velocity signals, and timings of ICs and FCs were determined from local maxima and minima in the heel and toe vertical velocity signals [PBv01; OCo+07]. All identified ICs and FCs were manually checked using Qualisys Track Manager 2018.1 software (Qualisys AB, Göteborg, Sweden) and corrected if necessary [Rom+21; Car+18]. The resulting annotated ICs and FCs were considered the true events (also

labels or targets), and were used as reference timings to derive stride-specific gait parameters.

IMU Data

The idea behind the deep learning approach was that a *model* was trained to predict the likelihood of an IC and FC, given accelerometer and gyroscope data from a single IMU. The data from a single sensor channel, e.g., the acceleration in forward direction, were denoted by $\mathbf{x}_d = \begin{bmatrix} x_d[1] & x_d[2] & \cdots & x_d[N] \end{bmatrix}^\mathsf{T}$, with d referring to the dth sensor channel (*i.e.*, $d = 1, 2, \cdots, D$) and n referring to the $n$the sample or time step (*i.e.*, $n = 1, 2, \cdots, N$). Similarly, the data from all D sensor channels at a given time step n were denoted by $\mathbf{x}[n] = \begin{bmatrix} x_1[n] & x_2[n] & \cdots & x_D[n] \end{bmatrix}^\mathsf{T}$. Data from all D channels, and for all N time steps, were then denoted by:

$$\mathbf{X} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_D \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} x_1[1] & x_2[1] & & x_D[1] \\ x_1[2] & x_2[2] & & x_D[2] \\ \vdots & \vdots & \cdots & \vdots \\ x_1[N] & x_2[N] & & x_D[N] \end{bmatrix}, \quad \mathbf{X} \in \mathbb{R}^{N \times D} \tag{4.1}$$

Likewise, the *labels* were denoted by:

$$\mathbf{y}_{\mathsf{IC}} = \begin{bmatrix} y_{\mathsf{IC}}[1] \\ y_{\mathsf{IC}}[2] \\ \vdots \\ y_{\mathsf{IC}}[N] \end{bmatrix}, \quad \mathbf{y}_{\mathsf{FC}} = \begin{bmatrix} y_{\mathsf{FC}}[1] \\ y_{\mathsf{FC}}[2] \\ \vdots \\ y_{\mathsf{FC}}[N] \end{bmatrix}, \quad \mathbf{y}_{\mathsf{IC/FC}} \in \mathbb{R}^{N \times 1}, \quad y_{\mathsf{IC/FC}}[n] \in \{0, 1\} \tag{4.2}$$

The model was iteratively trained to learn a mapping $h_{\mathbf{\Theta}}(\mathbf{X}) : \mathbf{X} \rightarrow \mathbf{y}$, where $h_{\mathbf{\Theta}}$ was also referred to as the *hypothesis* parameterized by the weights, collectively denoted by $\mathbf{\Theta}$, and $\mathbf{X}$ was an array with the raw sensor data from the 3-axis accelerometer and 3-axis gyroscope of a single sensor location.

All participant data were split into three independent datasets, namely a training set, a validation set, and a test set. Each set contained data from approximately one-third of the participants. Participants were randomly assigned to one of the sets, stratified by both group (i.e., diagnosis) and gender (Table 4.2). The training and validation set were used to train an optimal deep learning model. Test set data were not used for training the model or hyperparameter tuning. The results on the model's performance were only based on the test set, and therefore reflected how good the model generalizes to new, *unseen* data.

Table 4.2: Overview of the total number of participants, walking trials, and number of instances in the training, validation, and test set. A detailed overview with exactly for which trial and sensor location valid data were available can be found at https://github.com/rmndrs89/my-gait-events-tcn, accessed on 1 April 2022.

Dataset	Num. of participants	Num. of trials	Num. of instances
train	61	749	3366
validation	48	564	2570
test	48	620	620

Accelerometer and gyroscope data were normalized by subtracting the channel-wise mean and dividing by the channel-wise standard deviation. For the training and validation datasets, the data were then partitioned into equal length time windows [Fil+20] of 400 samples, with an overlap of $50\,\%$ between successive windows (corresponding to $2\,$s windows, and $1\,$s overlap, respectively). For the test set, the complete trial was fed as input to the model for predicting ICs and FCs (hence the number of instances is the same as the number of trials, Table 4.2).

4.2.3. Model

Model Architecture

The generic architecture for the deep learning model was based on a temporal convo- lutional network (TCN) [Fil+20; Rém20; BKK18]. The TCN consisted of a sequence of residual blocks with exponentially increasing dilation factor [BKK18; YK16]. Each residual block was built from two sequences of a dilated convolutional layer [YK16], a batch normalization layer [IS15], a rectified linear unit (ReLU) activation layer, and a dropout layer [Sri+14] (Figure 4.2). The model was built in Python [VD09] using the high-level TensorFlow API Keras [Cho+15; Rém20].

For the current analysis, only convolutions of the "same" type were considered [Rém20], *i.e.*, the model was non-causal and zero-padded to account for edge effects, and the likelihood of an IC or FC was based on input data both before and after the current sample, n:

$$\hat{y}_i[n] = f(\cdots, \mathbf{x}[n-1], \mathbf{x}[n], \mathbf{x}[n+1], \cdots), \quad i \in \{\mathrm{IC}, \mathrm{FC}\} \tag{4.3}$$

The number of samples that the predictions at time n "*sees*" was referred to as the receptive field [Col+17] and was a function of the kernel size and the dilation factors [Rém20]. Dilation factors were always given as a sequence of increasing power of 2 [BKK18; YK16; van+16a].

The outputs of the TCN block were fed to two separate fully connected (FCN, or dense) layers, that were both followed by a sigmoid activation layer. Outputs were then predicted separately for

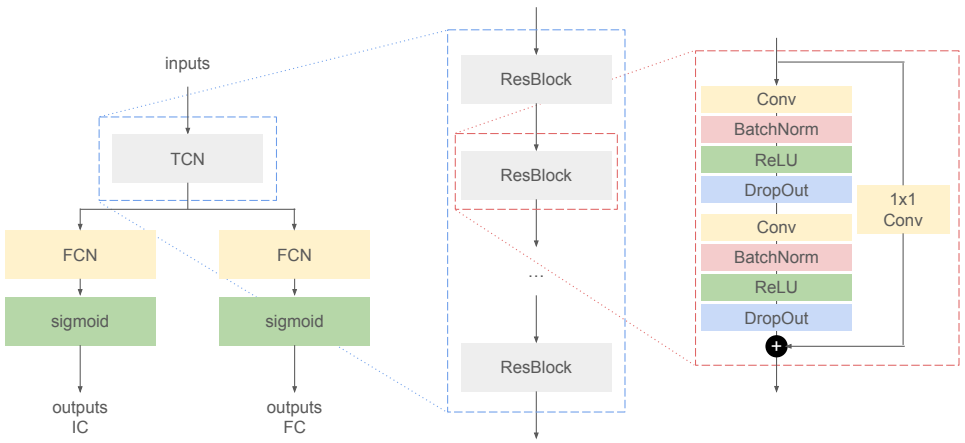


Figure 4.2: The generic model architecture of the deep learning model to predict initial contacts(ICs) and final contacts (FCs). The inputs are the accelerometer and gyroscope data from a single inertial measurement unit, which are fed to a temporal convolutional network (TCN) (**left**). The TCN consisted of repeating residual blocks (ResBlocks) with exponentially increasing dilation factor (**middle**). Each ResBlock was built from two sequences of a convolutional layer (Conv), batch normalization layer (BatchNorm), a rectified linear unit activation layer (ReLU), and a dropout layer (DropOut) (**right**).

ICs and FCs [Fil+20; Gad+19]. The mean squared error (MSE) was used as a loss function, and a gradient descent-based optimization algorithm with adaptive moment (Adam) optimizer was used to iteratively learn the weights [KB14; SSH21].

Hyperparameter Optimization

In order to find the best model architecture, hyperparameter tuning was performed using KerasTuner [OMa+19]. Here, the number of filters, the kernel size, and the maximum dilation factor (Table 4.3) were optimized for using a random search strategy [BB12].

Table 4.3: Model hyperparameters that were optimized for, and the corresponding sets of possible values.

Description	**Possible values**
number of filters	8, **16**, 32, 64, 128
kernel size	3, **5**, 7
dilation	[1, 2], **[1, 2, 4]**, [1, 2, 4, 8]

The model architecture that resulted from the hyperparameter optimization was then trained on

the combined set of training and validation data. The trained model was used to predict the likelihoods of ICs and FCs on the test set data.

4.2.4. Analysis

The predictions of the model on the test set data were compared with the labels from the test set. The model performance was evaluated for (1) overall detection performance, (2) time agreement between the predicted events and the (marker-based) annotated events, and (3) agreement between subsequently derived stride-specific gait parameters.

Overall Detection Performance

The overall detection performance quantified how many of the annotated events were detected by the model (true positives, TP), how many of the annotated events were not detected (false negatives, FN), and how many events that were detected, were actually not annotated (false positives, FP). From these metrics, the recall (or sensitivity), precision, and F1 score were calculated as:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4.4}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4.5}$$

$$\text{F1 score} = 2 \cdot \frac{\text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} \tag{4.6}$$

Here, recall represented the ratio of gait events that were detected, precision represented the ratio of detected gait events that were truly gait events, and F1 score was the harmonic mean of the recall and precision.

Time Agreement

For all correctly detected gait events (TP, Section 4.2.4), time agreement was assessed by the time error between the annotated and detect gait event, which was defined as:

$$\text{time error} = t_{\text{ref}} - t_{\text{pred}} \tag{4.7}$$

with t_{ref} the gait event time from the marker-based annotations, and t_{pred} the gait event time from the model predictions. As a robust measure for the average time error and its spread, the median time error and the inter-quartile range (IQR) were reported [DÇB19].

Stride-Specific Gait Parameters

For those trials for which all gait events were detected and no false positives were detected, the stride time, stance time, and swing were calculated. Stride was the time between two successive ICs of the same foot. Stance time was the time between a FC and the preceding IC of the same foot. Swing time was the time between the IC following the last FC of the same foot.

4.3. Results

4.3.1. Overall Detection Performance

The performance of detecting ICs and FCs was objectively quantified by the number of TPs, the number of FNs, and the number of FPs. From these numbers, recall, precision, and F1 score were calculated (Table 4.4).

Table 4.4: Overall detection performance for initial contacts and final contacts as quantified by recall (R), precision (P), and F1 score. TP: true positives, FN: false negatives, FP: false positives.

	Initial contacts						**Final contacts**					
Tracked point	**TP**	**FN**	**FP**	**R**	**P**	**F1**	**TP**	**FN**	**FP**	**R**	**P**	**F1**
left ankle	624	19	5	$97\,\%$	$99\,\%$	$98\,\%$	606	32	10	$95\,\%$	$98\,\%$	$97\,\%$
right ankle	599	42	8	$93\,\%$	$99\,\%$	$96\,\%$	614	17	12	$97\,\%$	$98\,\%$	$98\,\%$
left shank	605	38	15	$94\,\%$	$98\,\%$	$96\,\%$	585	53	18	$92\,\%$	$97\,\%$	$94\,\%$
right shank	603	36	15	$94\,\%$	$98\,\%$	$96\,\%$	595	30	9	$95\,\%$	$99\,\%$	$97\,\%$

For both ICs and FCs, recall is high for each of the sensor locations (i.e., $\geq 92\,\%$) and so is precision (i.e., $\geq 97\,\%$). Differences between the sensor locations are small, i.e., the minimum recall is $92\,\%$ and the maximum recall is $97\,\%$, and the minimum precision is $97\,\%$ and the maximum precision is $99\,\%$. The recall and precision result in F1 scores of $\geq 96\,\%$ for ICs and $\geq 94\,\%$ for FCs.

4.3.2. Time Agreement

Time agreement between the annotated and detected events was quantified for the TPs for each of the sensor locations (Figure 4.3). The median time error for each of the sensor locations and for both ICs and FCs was close to zero (Table 4.5), and the largest median time error was $-0.005\,\mathrm{s}$, corresponding to one sample period (at a sampling frequency of $200\,\mathrm{Hz}$). The IQR was at most $0.020\,\mathrm{s}$, corresponding to four sample periods.

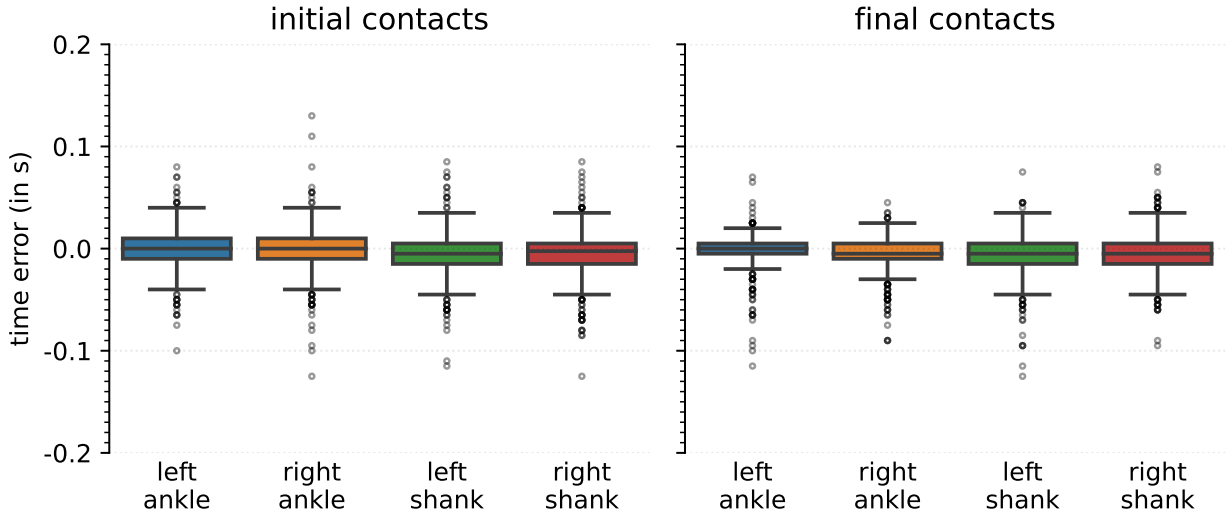


Figure 4.3: Time errors for the initial (**left**) and final (**right**) contacts detection, for each of the different tracked points.

Table 4.5: Time errors for the correctly detected gait events. Note that 5 millis corresponds to 1 sample period, given the sampling frequency of 200 Hz.

	Initial contacts		**Final contacts**	
Tracked point	**median**	**IQR**	**median**	**IQR**
	ms	ms	ms	ms
left ankle	0	20	0	10
left shank	0	20	-5	15
right ankle	-5	20	-5	20
left ankle	-3	20	-5	20

4.3.3. Stride-Specific Gait Parameters

For those trials for which all gait events were correctly detected (and no false positives were detected), stride time, stance time, and swing time were calculated. The mean difference and the limits of agreement between the marker-based annotations and the model-based detections were calculated. For all stride-specific gait parameters, and for all sensor locations, the mean difference was close to zero, i.e., the maximum mean difference was 0.003 s, namely for the calculated swing time of the right ankle (Table 4.6). Furthermore, for all gait parameters and for all sensor locations, the limits of agreement, based on a 95 % confidence interval, were distributed around a zero-mean

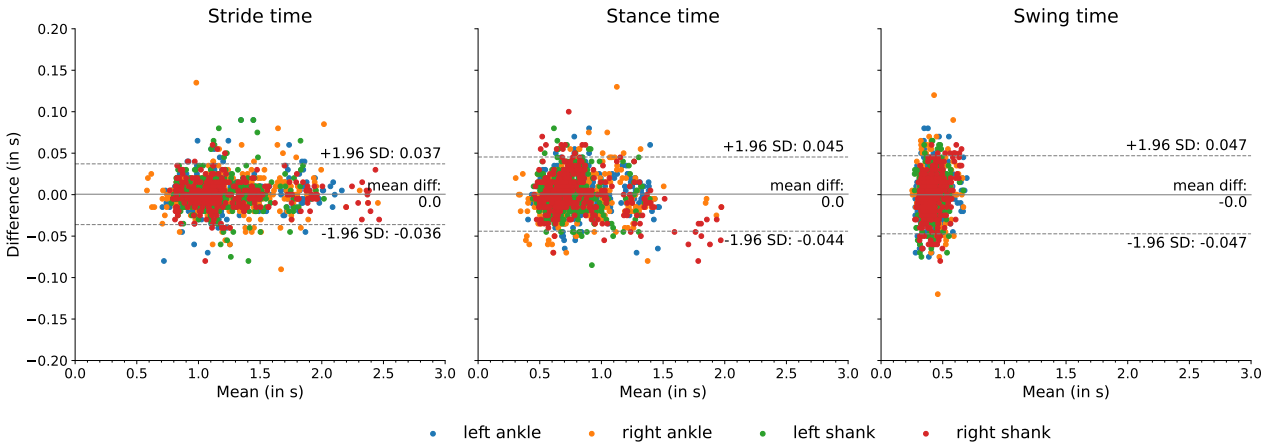


Figure 4.4: Time errors for the initial (**left**) and final (**right**) contacts detection, for each of the different tracked points.

difference with the overall limits of agreement at -0.049 s and 0.051 s (Figure 4.4).

Table 4.6: Time agreement between the stride-specific parameters.

Tracked point	**Parameter**	**Mean difference**	**Limits of agreement**
		s	(s, s)
left ankle	stride time	0.001	(−0.035, 0.036)
	stance time	0.002	(−0.039, 0.042)
	swing time	−0.001	(−0.045, 0.043)
right ankle	stride time	0.000	(−0.039, 0.041)
	stance time	−0.002	(−0.048, 0.044)
	swing time	0.003	(−0.046, 0.051)
left shank	stride time	0.001	(−0.039, 0.041)
	stance time	0.002	(−0.043, 0.046)
	swing time	−0.001	(−0.049, 0.047)
right shank	stride time	−0.000	(−0.031, 0.031)
	stance time	0.002	(−0.046, 0.049)
	swing time	−0.002	(−0.049, 0.046)

4.4. Discussion

The current study aimed to validate a deep learning approach for detecting gait events from a single IMU worn on the lower leg. Data from left and right ankle- and shank-worn IMUs were used for

training a neural network to detect gait events from walking trials performed by healthy YA, healthy OA, participants diagnosed with PD, MS, or cLBP, participants who had a recent symptomatic stroke, and participants diagnosed with other neurological diseases. Participants walked a 5 m distance at three different self-selected walking speeds. The gait event timings that were predicted by the neural network were compared to a common reference method, i.e., OMC system, and clinically relevant stride-specific gait parameters were extracted.

A first measure for the model performance was given by the recall (how many annotated events were detected) and precision (how many detected events were annotated). For both ICs and FCs, a high recall (≥ 95 %), high precision (≥ 95 %), and high F1 score (≥ 94 %) were observed, meaning that most events were detected and most detected events were actually true events. There was little difference in recall, precision, and F1 score between sensor locations (Table 4.4), confirming that the deep learning approach is relatively invariant to exact sensor localization. The values for recall, precision, and F1 score are comparable to recall (≥ 85 %), precision (≥ 95 %), F1 score (≥ 91 %) from studies that detected gait events in OA, people with PD, and people with MS [Rom+21] or adults and hemiplegic patients [Ji+19].

Next, the time error, that is, the difference between the annotated event and the detected event, was of interest. For both ICs and FCs, and for all sensor locations, the observed time error was small, and the middle 50 % of the time errors were within a range of $[-0.015, 0.010]$ s (Table 4.5, Figure 4.3). These data showed that the deep learning-based approach is precise in detecting initial and final contacts. Time errors were slightly smaller than our previously reported results [Rom+21] that used a heuristics-based approach [Sal+04]. The heuristics-based approach determined ICs and FCs as local minima in the medio-lateral angular velocity [Sal+04], and it could be that these minima do not exactly coincide with the reference event timing as determined from the OMC systems. Previous studies that investigated the time error of IMU-based gait event detection reported a 95 % confidence interval of $[0.007, 0.013]$ s for IC, and $[-0.005, 0.004]$ s for FC for young and older adults in treadmill and overground walking [Ami+02], or $[-0.016, 0.001]$ s for IC or $[0.037, 0.063]$ s for FC for typically developing children in overground walking [CGE10]. Others reported the mean for the time error in healthy elderly subjects, subjects with PD, subjects with choreatic movement disorder, and hemiparetic subjects, and found a maximum mean error of 0.011 s at normal walking speed, and 0.022 s at faster speed [Tro+14]. Similarly, for healthy subjects, a mean error of 0.017 s for ICs and -0.016 s for FCs were reported, whereas for a single transfemoral amputee the mean error was 0.012 s for ICs and -0.024 s for FCs for the intact limb [Maq+16]. The median and IQR of the time errors of the current study were in the same range as previously found by [Gad+19], and time errors were smaller than previously reported time errors from a continuous wavelet-based approach [Ji+19]. Hence, our proposed deep learning approach resulted in time errors that are in the same range or better than those from previous approaches, while not being restricted to an exact sensor location

(left or right ankle, or shank) and sensor alignment.

From the correctly detected gait events, stride-specific gait parameters were derived. These are probably of greatest clinical relevance, as changes in stride-specific gait parameters have been linked directly with disease onset and progression [Del+19; Kön+17; Ber+18; von+95; Moh+21; Gri+16; Fla+19]. Therefore, stride time, stance time, and swing time were calculated, and the differences between the deep learning-based approach and the marker-based reference method were quantified (Table 4.6, Figure 4.4). The limits of agreement for a 95 % confidence interval were calculated, and for all metrics the zero-mean difference was enclosed in the limit of agreement. These data provided evidence that the deep learning-based model was able to derive stride-specific gait parameters. The differences between the deep learning model-based stride parameters and the marker-based stride parameters were in a similar range as a recent study that compared IMU-derived stride parameters against stride parameters obtained with a pressure sensing walkway [NK21; Gad+19] and were also in the same range as results from a study that compared IMU-derived stride parameters with stride parameters obtained with an OMC systems [Car+18]. The mean error was lower than the mean errors reported for stance and swing time (0.011 s and 0.011 s, respectively) across elderly subjects, subjects with PD, subjects with choreatic movement disorder, and hemiparetic subjects [Tro+14].

The main limitations of the current study were that only walking trials involving straight-line walking were considered, and the walking distance was relatively short. Therefore, it may be that the observed gait patterns from these walking trials are not fully representative of gait patterns observed in daily life [Hil+19; War+20; Atr+21]. However, as the deep learning-based approach does not rely on fixed thresholds or assumptions of which sensor axis is used, it is theoretically transferable and scalable to other conditions if input data and corresponding labels can be provided.

Furthermore, although the proposed approach allows for relatively arbitrary sensor placement on the lower leg, it was not investigated to what extent participants would be willing to wear such a sensor for a prolonged period of time in the home environment. Previous research found that user acceptance and adherence to wearing IMUs was generally high in people with neurodegenerative diseases [van+16b; Fis+16; Ada+17; God+22], although reduced adherence was linked with multi-day wear [Fis+16] and wearing multiple sensors [Esp+19].

4.5. Conclusion

In this study we have validated a DL-based approach to detect gait events and subsequently extract clinically relevant stride-specific gait parameters from a single IMU worn either laterally above the ankle joint or proximal below the knee joint. Performance analysis showed an excellent detection rate and low time errors in both event detection and stride parameter calculation for different walking

speeds and across both healthy and neurological cohorts. Compared to relevant approaches that detected gait events from an ankle- or shank-worn IMU, the DL approach reached a performance that was on par or better, and it did not rely on expert-defined, hand-crafted features or empirically derived thresholds. The performance of the DL approach was not affected by the exact sensor placement and orientation, and hence is less obtrusive for potential applications in long-term continuous monitoring. In contrast to previous approaches, it allows for personalization of the network to individual study participants and is easily transferable even to other sensor placement locations (e.g., a foot-worn or low back-worn IMU) without the need for rethinking the set of decision rules and thresholds. Our next step is to further develop and validate these methods with real-life walking sequences in patients with neurodegenerative diseases.

Acknowledgements

The authors sincerely thank all people that were involved in the data collection, from study participants to students who aided and assisted during measurements and subsequent marker labeling. The authors thank Julius Welzel for his input regarding the organization of the data in a BIDS-like format. The authors thank Johannes Hoffmann for his tips regarding the LaTeXtypesetting.

Funding

We acknowledge financial support by Land Schleswig-Holstein within the funding programme Open Access Publikationsfonds.

Contributions

Conceptualization, Robbin Romijnders (R.R.), Gerhard Schmidt (G.S.) and Walter Maetzler (W.M.); methodology, R.R. and G.S.; software, R.R.; validation, R.R.; formal analysis, R.R.; investigation, R.R., G.S. and W.M.; resources, Clint Hansen (C.H.) and W.M.; data curation, Elke Warmerdam (E.W.) and R.R.; writing—original draft preparation, R.R.; writing—review and editing, E.W., C.H., G.S. and W.M.; visualization, R.R.; supervision, G.S. and W.M.; project administration, R.R., E.W., C.H., G.S. and W.M.; funding acquisition, W.M. All authors have read and agreed to the published version of the manuscript.

Ethics approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the medical faculty of the Christian-Albrechts-Universität zu Kiel (D438/18, approved on 8 May 2018).

5

Ecological Validity of Deep Learning for Gait Events Detection

The contents of this chapter are based on:

Romijnders, R., Salis, F., Hansen, C., ..., Schmidt, G., Maetzler, W., & Mobilise-D consortium (2022). Ecological validity of a deep learning algorithm to detect gait events from real-life walking bouts in mobility-limiting diseases. *Submitted to Digital Health.*

Abstract

Background

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, extraction of clinically relevant gait parameters from IMU data often depends on heuristics-based algorithms that rely on empirically determined thresholds that were mainly validated on small cohorts in supervised settings. Here, a deep learning (DL) algorithm was developed and validated for gait events detection in a heterogeneous population of different mobility-limiting disease cohorts and a cohort of healthy adults.

Methods

Participants wore pressure insoles and IMUs on both feet for 2.5 hours 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

The results show 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 s, -0.07 s and <0.01 s for stance, swing and stride time, respectively.

Conclusions

Thus, the DL algorithm is considered successful in detecting gait events in ecologically valid environments across different mobility-limiting disease.

5.1. Introduction

Mobility is the ability to move about in the home and community [Roc+20]. Mobility can be affected by chronic health conditions, including but not limited to neurological, respiratory, cardiac and musculoskeletal disorders [Klu+21]. Deficits in mobility have been linked with a reduced quality of life, an increased fall risk, and mortality [Klu+21; War+20], therefore mobility is regarded as an essential aspect of health [PR91]. The most common and functionally relevant aspect of mobility that is affected by ageing and chronic health conditions is walking [Roc+20; Sni+07].

To date, clinical assessment of mobility is based on functional tests that include short walking tasks [But+82; Gur+94; Mot+17; PR91]. 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 [Atr+21; FM20; Hil+19; War+20]. The transition to unsupervised monitoring of human motion in naturalistic and unconstrained daily-life activities is driven mainly by the use of wearable inertial measurement units (IMUs) [Pic+21; PR91]. 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 [Pic+21; Rav+19]. 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 [Cor+19; FM20].

A critical step for the objective analysis of gait is the segmentation of gait sequences into gait cycles [Cel+21; Bob+18], i.e., the basic repetitive unit that gait is comprised of [PB10; LRW12]. 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 [PB10; LRW12]. Together with the instant at which the foot leaves the ground (i.e., final contact, FC), each stride can be divided in a stance and swing phase [PB10; LRW12; Rue+10; BR14]. The detection of ICs and FCs from IMUs is typically done using heuristics-based algorithms [Ami+02; Sal+04; Par+04; Tro+14; Par+19; PSA20; Rom+21]. Many of these algorithms use local maxima or minima of the acceleration and/or angular velocity signals along one axis [NK21] which requires knowledge of the sensor-to-segment alignment [Pac+20; LGA19]. However, in unsupervised human gait monitoring, the sensor-to-segment alignment cannot be controlled for as study participants often attach the sensor themselves, for example, after showering [Rom+22]. 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 labelled free-living gait data [Cam+18; HHP16; Ray+21]. Additionally, IMU-based gait signals are affected by disease characteristics, participant activity levels and the exact context in which walking takes place, and therefore any heuristics-based algorithm that was developed based on lab-based gait data might not translate

directly to free-living gait [Pan+18; Cor+19; FM20; Rom+21; War+20].

In contrast to the afore-mentioned 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 [Haj+18; Gér19; Cho21]. Hidden Markov models (HMMs), for example, were successfully applied for gait segmentation in healthy [MS11; Pan+13] and pathological gait [MS11; Mar+17], but only in-lab recorded gait data was 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 was only from participants with Parkinson's disease (PD) [Rot+21c]. Although, HMMs thus seem a good fit for modelling 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 [Mar+21; Rot+21a]. 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 [LBH15; Esk+16; Gér19; Cho21]. DL algorithms have been successfully applied for gait events detection from stereophotogrammetric data [Chi+05; Mil09; KDS19; Fil+20; Lem+20] and from inertial measurement units data [Gad+19; Rom+22], however, only for in-lab gait data.

Therefore, the specific aim of the current study was to determine if a previously in-lab validated DL-based gait event detection algorithm [Rom+22] can be used for the detection of gait events from real-life walking bouts in a heterogeneous cohorts of different mobility-limiting diseases.

5.2. Methods

5.2.1. Data Collection

Study Participants

As part of the Mobilise-D technical validation study [Maz+21], a convenience sample of 108 participants were recruited at five independent study sites ((1) Newcastle upon Tyne Hospitals NHS Foundation Trust, UK; (2) Sheffield Teaching Hospitals NHS Foundation Trust, UK; (3) Tel Aviv Sourasky Medical Center, Israel; (4) Robert Bosch Foundation for Medical Research, Germany; (5) University of Kiel, Germany). The sample represented five mobility-limiting disease cohorts ((1) congestive heart failure (CHF), (2) chronic obstructive pulmonary disease (COPD), (3) multiple sclerosis (MS), (4) Parkinson's disease (PD), (5) proximal femoral fracture (PFF)) and a cohort of healthy older adults (HA) [Maz+21].

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 [Ber+17; BDC19; Sal+19; Tro+14]. Participants wore the INDIP

system for about 2.5 hours in their habitual environment, e.g., home, work, community, and/or outdoor environment, which was chosen by the participant, with no specific restrictions [Maz+21]. To capture the largest possible range of activities, participants were provided a list of activities that could be included, if relevant for their chosen environment (e.g., rise from a chair, walk to another room, walk outdoors). No supervision or structure to how these tasks were completed was given to the participants.

5.2.2. Data Processing

Data Preparation

Data was split into three different data sets: a training set, a validation set, and a testing set [Gér19; Cho21]. 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 [Ped+11], and the training set was used to optimize the corresponding model parameters [Gér19; Cho21]. The testing set was only used for the final evaluation, and notably the numbers presented in Results section only corresponded to the performance on the testing set.

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. Then, results were combined, and priority was given to the PIs in case both modalities detected an event [Sal+21b]. For the PIs, foot-ground contact was defined when at least three sensing elements from the PI belonging to the same spatial neighbourhood were consecutively activated and deactivated [Sal+21a]. 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 detection of supervised gait events in older, hemiparetic, parkinsonian and choreic gait [Tro+14; Ros+21] and across multiple research centers for parkinsonian and mildly cognitive impaired gait [Ber+18].

From these gait events, walking bouts (WBs) were formed by merging information from left and right strides. Each WB represented a gait sequence with a minimum of two left and two right strides [Klu+21; Sal+21b]. Here, strides were only considered valid if (i) the stride duration was between $0.2\,$s and $3\,$s, and (ii) the stride 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 [Bru+09; Ris+15; van+14; van+15] and for which both the INDIP's PIs and IMUs were used for determining the gait events. These gait events were considered as the reference annotations for training and evaluating the DL

algorithm.

Deep Learning-Based Algorithm

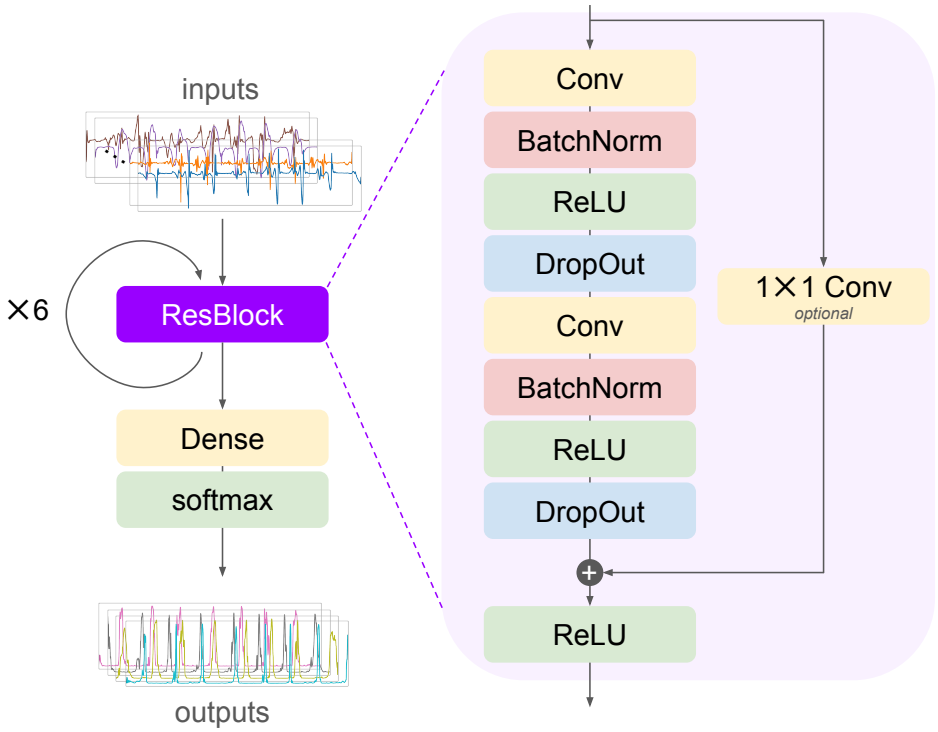


Figure 5.1: Schematic depiction of the deep learning model architecture with a residual block (ResBlock) that is repeated (in this case 6 times) before a dense and softmax layer are applied. Inputs to the network are the raw accelerometer and gyroscope data of both left and right inertial measurement unit. The outputs are estimated probabilities for each of the gait events for each time step. Conv: Convolution, BatchNorm: batch normalization, ReLU: recitified linear unit, DropOut: dropout

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 [War+21; Rom+22]. At the core of the NN was a temporal convolutional network (TCN) [BKK18; Rém20]. The TCN was built from stacking residual blocks [He+16] with an exponentially increasing dilation factor for the convolutional layers (Figure 5.1).

Specifically, each residual block comprised two sequences of a dilated convolution (Conv) layer [YK16; YKF17], a batch normalization (BatchNorm) layer [IS15], a rectified linear unit (ReLU) activation layer, and a dropout layer [Sri+14]. A residual connection was used that performed convolution with a kernel size of 1 in case the number of feature maps did not match the number of input channels [BKK18; Rém20]. The outputs of the second dropout layer and the residual connection were

summed element-wise, and inputted to a ReLU activitation layer. The convolution layers consisted of 64 filters with a kernel size of 3 and a dilation factor of 2^{m-1} with $m = 1, 2, \cdots, N_{\text{dil}}$ for the m-th residual block ($N_{\text{dil}} = 6$, 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 [Bis06; GBC16]. The final outputs were then regarded as probabilities that a certain gait event took place at the given time step, t_n.

5.2.3. Evaluation

As in our previous works [Rom+22], the performance was evaluated with the testing set data only. The trained model was used to predict from the IMU data the probability that any gait event occurred. Peak probabilities, with a minimum probability, $\Delta_{\text{Pr}} = 0.5$, and a minimum interpeak distance, $\Delta_{\Delta t} = 0.5\,\text{s}$, were considered detected events.

Performance was evaluated for (1) overall detection performance, (2) time agreement between predicted and annotated gait event timings, and (3) time agreement between subsequently derived stride-specific gait parameters.

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 a sensitivity) and precision (also referred to as positive predictive value) were calculated as:

$$\text{recall} = \frac{\#\,\text{true positives}}{\#\,\text{true positives} + \#\,\text{false negatives}} \tag{5.1}$$

$$\text{precision} = \frac{\#\,\text{true positives}}{\#\,\text{true positives} + \#\,\text{false positives}} \tag{5.2}$$

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 $\leq 250\,\text{ms}$ it was considered a true positive event [Pha+17; Rom+21; Bon+22; Rom+22] (in other words, a tolerance window of $500\,\text{ms}$ centered around the reference timing was used).

Time Agreement

For all correctly detected gait events (true positives, Section 5.2.3) the time agreement between the detected and annotated event timings was quantified by:

$$\text{time error} = t_{\text{ref}} - t_{\text{pred}} \tag{5.3}$$

with t_{pred} the timing corresponding to the peak probability, and t_{ref} the timing from the INDIP-derived annotations.

As a robust measure for the time agreement and its spread, the median time error and the interquartile range (IQR) were computed [DÇB19], and time agreements were visualized using box plots.

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 [PB10; LRW12; DGR16]. Stance time was the time between a 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 [DGR16; Rom+22].

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 [DÇB19]. Differences were visualized using Bland-Altman plots [AB83; Gia15].

5.3. Results

5.3.1. Demographics

Data from 99 participants were used for the current study (Table 5.1). Data from the other participants were excluded due to incomplete or missing data of the INDIP system, or because no WBs $\geq 10\,\text{s}$ were recorded. Eventually the DL-based algorithm was evaluated for its performance in detecting gait events of 13100 strides divided over 295 bouts recorded from 17 participants in the testing set.

Table 5.1: Demographics data of the study participants summarized by set and cohort.

Set	Cohort	Num. of participants	Age	Height	Mass	Num. of bouts	Num. of strides
train	CHF	8	69 (13)	177 (8)	86 (20)	189	11326
	COPD	11	70 (9)	169 (6)	73 (14)	187	6562
	MS	12	47 (8)	171 (14)	80 (23)	139	6216
	PD	12	70 (7)	175 (6)	79 (16)	165	7574
	PFF	10	83 (6)	172 (9)	71 (16)	151	5838
	HA	12	71 (7)	168 (10)	76 (11)	245	13597
validation	CHF	2	74 (13)	172 (21)	87 (3)	41	1210
	COPD	3	69 (14)	171 (10)	69 (12)	68	1890
	MS	3	42 (15)	172 (13)	97 (24)	24	863
	PD	3	70 (7)	174 (6)	79 (21)	61	3466

Table 5.1: Demographics data of the study participants summarized by set and cohort.

Set	Cohort	Num. of participants	Age	Height	Mass	Num. of bouts	Num. of strides
	PFF	2	71 (1)	164 (8)	60 (9)	31	1078
	HA	4	72 (4)	169 (10)	77 (18)	216	4952
test	CHF	2	65 (13)	168 (1)	77 (16)	10	407
	COPD	3	69 (8)	166 (3)	80 (18)	79	2346
	MS	3	58 (12)	172 (16)	87 (25)	53	2576
	PD	3	70 (11)	166 (11)	73 (8)	38	2448
	PFF	2	76 (6)	168 (8)	75 (28)	21	1649
	HA	4	73 (3)	164 (11)	72 (10)	94	3674

5.3.2. Overall Detection Performance

Table 5.2: Overall detection performance of initial and final contacts evaluated per cohort. CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, FN: false negative, FP: false positive, HA: healthy adults, MS: multiple sclerosis, PD: Parkinson's disease, PFF: proximal femoral fracture, TP: true positive.

	Initial contacts					Final contacts				
	TP	FN	FP	R	P	TP	FN	FP	R	P
CHF	408	3	18	99%	96%	401	1	23	100%	95%
COPD	2294	58	86	98%	96%	2235	37	147	99%	94%
MS	2563	19	72	99%	97%	2518	12	95	100%	96%
PD	2431	23	40	99%	98%	2411	3	55	100%	98%
PFF	1642	11	15	99%	99%	1614	18	45	99%	97%
HA	3627	55	98	99%	97%	3568	20	141	99%	96%

The overall detection performance was quantified by the number of TPs, number of FNs, and number of FPs. From these numbers the recall and precision were calculated (Table 5.2). In total, from 13134 ICs, the algorithm detected 12985 events (i.e., 99%) and missed 169 events (i.e., 1%), and likewise from 12838 FCs, the algorithm detected 12747 events (i.e., 99%) and missed 91 events (i.e., 1%). When evaluated per cohort, the recall for the ICs detection was $\geq$ 98% and the precision was $\geq$ 96%. Similarly, the recall was $\geq$ 99% and the precision was $\geq$ 94% for FCs detection for all

cohorts.

5.3.3. Time Agreement

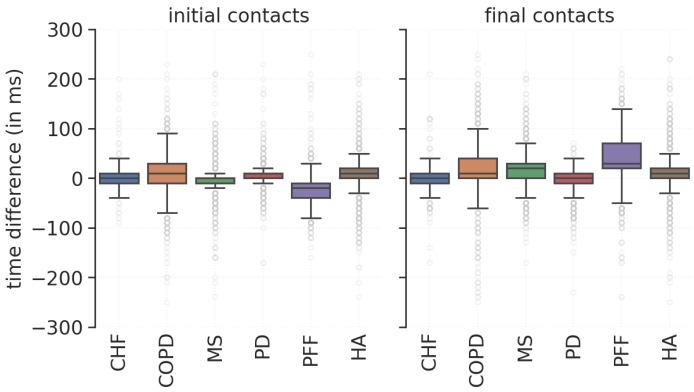


Figure 5.2: Time difference between the predicted and reference events timings for initial and final contacts evaluated per cohort. A positive time difference corresponded to an advanced detection. CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, MS: multiple sclerosis, HA: healthy adults, PD: Parkinson's disease, PFF: proximal femoral fracture

For all the correctly detected events, that is the TP events, the difference between the detected event timing and the annotated event timings was calculated according to Equation 5.3. 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 5.2). The PFF cohort showed a median time error of -0.02 s and an IQR of 0.03 s for ICs detection, and a median time error of 0.03 s and IQR of 0.05 s for FCs detection (Table 5.3).

Table 5.3: Time differences between the predicted event timings and the annotated event timings evaluated per cohort.

	Initial contacts		**Final contacts**	
Cohort	**median**	**IQR**	**median**	**IQR**
	ms	ms	ms	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

5.3.4. 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 5.3). Notably, for the PFF cohort the mean time difference for the stance time was $+0.07$ s, the mean time difference for the swing time was -0.07 s, which resulted in a zero-mean difference for the stride time (Table 5.4). Similarly, for all gait phases the absolute errors were 0.04 s or less for all cohorts, except the PFF cohort (Table 5.5). This resulted in 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 12 %.

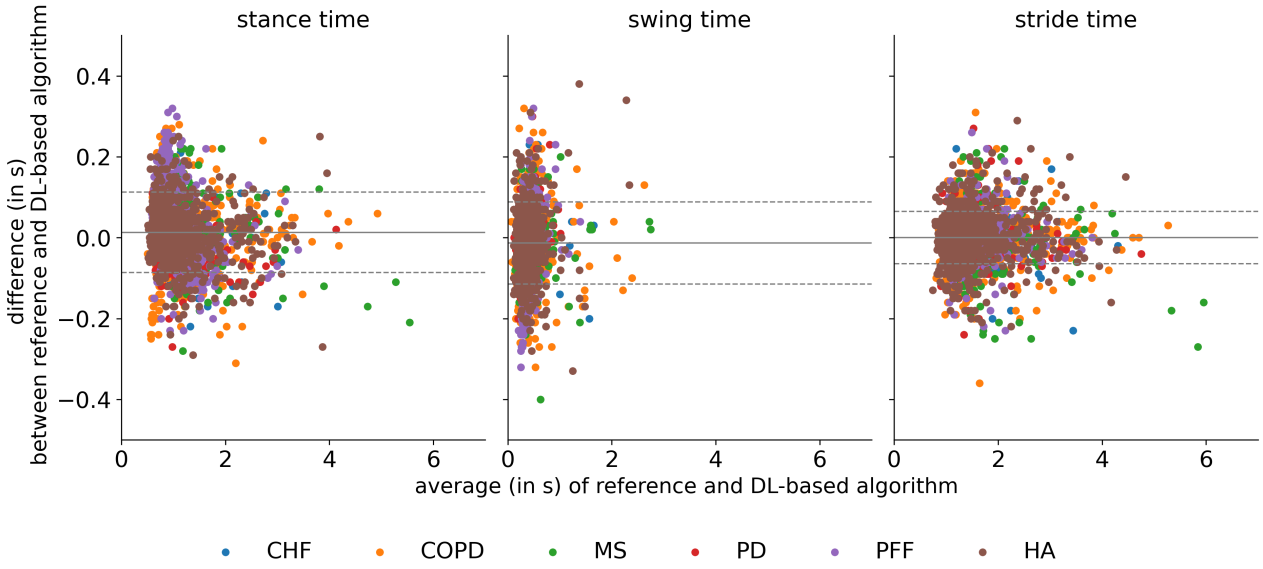


Figure 5.3: Bland-Altman plots [AB83] for the stance, swing and stride times evaluated per cohort. The gray solid line corresponded to the overall mean difference, and the dashed lines corresponded to the mean difference $\pm$ 1 standard deviation. CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, DL: deep learning, HA: healthy adults, MS: multiple sclerosis, PD: Parkinson's disease, PFF: proximal femoral fracture

Table 5.4: 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	**LoA**	**Mean difference**	**LoA**	**Mean difference**	**LoA**
	ms	(ms, ms)	ms	(ms, ms)	ms	(ms, ms)
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)

Table 5.5: Stance, swing and stride times obtained from the reference and the DL algorithm, and the absolute and relative time errors for comparison. Values represent the mean and 95% confidence interval of all stances, swings, and strides of the test subjects for the given cohort.

Gait phase	Cohort	Reference		DL algorithm		Absolute time error		Relative time 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)
	HA	0.84	(0.83, 0.85)	0.84	(0.83, 0.85)	0.03	(0.02, 0.03)	3	(3, 3)
	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)
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)
	HA	0.36	(0.36, 0.36)	0.36	(0.36, 0.37)	0.03	(0.03, 0.03)	8	(8, 9)
	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)
stride	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)
	HA	1.20	(1.19, 1.21)	1.20	(1.19, 1.21)	0.02	(0.02, 0.02)	2	(1, 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)

CHF: congestive heart failure, COPD: chronic obstructive pulmonary disease, HA: healthy adults,

MS: multiple sclerosis, PD: Parkinson's disease, PFF: proximal femoral fracture.

5.4. Discussion

The specific aim of the current study was to determine if a previously in-lab validated DL-based gait event detection algorithm [Rom+22] could be used for the detection of gait events from real-life walking bouts in a heterogeneous cohort of different mobility-limiting diseases. For that purpose, participants from different disease cohorts (CHF, COPD, MS, PD, 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 about 2.5 hrs 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 input to the DL algorithm to identify ICs and FCs.

Data were collected from 108 different participants, and eventually data from 99 participants could be included in the final analysis. These participants were divided in a training, validation and testing data set. The DL algorithm's performance was evaluated only for a hold-out testing set that consisted of 17 participants (Table 5.1). The INDIP reference system registered 295 walking bouts with a minimum duration of ≥ 10 s and 13100 strides. When evaluated per cohort, the recall was $\geq$ 98% and precision was $\geq$ 96% for ICs detection, and the recall was $\geq$ 99% and precision $\geq$ 94% for FCs detection, with the lowest precision (94%) observed for the FCs detection of the PFF cohort. For comparison, in [Tro+14] no missed or extra gait events were observed in a heterogeneous sample of elderly, hemiparetic, parkinsonian and choreic gait, but data was only collected from walking back and forth for about one minute in a 12 m walkway. Likewise, high recall and precision ($\geq$ 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 [Ji+19]. A recent study [Rot+21c] found a recall of 96% and precision of 89% in a cohort of 28 PD participants, who wore two IMUs on the feet for two 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 per cohort by the median time difference and the IQR. For all cohorts, except the PFF cohort, the median time difference was close to 0 s and the middle 50% of the time differences were in the range between -0.04 s and 0.01 s. For the PFF cohort, the median time difference for ICs detection was 0.02 s with the middle 50% of the time differences ranging from 0.01 s to 0.04 s, and the median time difference for FCs detection was -0.03 s with the

middle 50% of the time differences ranging from -0.07 s to -0.02 s. These results suggested that IC detection tended to be a bit delayed, whereas FCs detection tended to be slightly anticipated. However, the time differences were still in the same range as those previously reported for CWT-based [Ami+02; CGE10; Tro+14; Ji+19; Rom+21] and DL-based algorithms [Rom+22] 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 [Cat+08; Bob+18; Sal+21a]. In particular, for the INDIP pressure insole method [Sal+21a] while a negligible delay (<10 ms) was observed for FCs, a consistent ICs anticipation (20 ms) was found when compared to force plates [Sal+21a]. 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 have been linked with conversion to PD [Del+19] and from mild cognitive impairment to Alzheimer's disease [Kön+17], and values of temporal gait parameters were different in disease cohorts compared to healthy cohorts [Gri+16; Fla+19; Moh+21]. Here, a zero-mean time difference was found for the stride times for all cohorts. Likewise, 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 [Gau+18; SRT22]. Nonetheless, the time agreement for the stride-specific temporal gait parameters derived from the DL algorithm and the reference system were in a similar range as those communicated before for a comparable DL-based approach that evaluated results only from straight-line walking in supervised laboratory setting [Gad+19].

The very good results that were obtained in the current study for two feet-worn IMUs [Maz+21] combined with the results for a single shank-worn IMU from our previous study [Rom+22] provided evidence that the algorithm performance generalizes to other sensor wear locations and to free-living gait data. Furthermore, the algorithm performance was evaluated across 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.

Limitations of the current study included that only data from detected WBs were used. This means that gait events detection relied on the accurate detection of gait sequences as a preceding step [Rot+21c]. However, several algorithms have been reported for accurate IMU-based gait sequence detection in both healthy and disease cohorts [Par+04; Sal+04; Kar+06; DKZ10; Ilu+14;

Par+19; PSA20; Ull+20; Mar+21]. Furthermore, data from some participants had to be excluded from analysis due to missing or incomplete data (Section 5.3.1) 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 [BR14], it showed the difficulty of obtaining a data set with annotated gait events on completely unsupervised free-living gait data [HHP16; Cam+18; Ray+21; Rot+21c].

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, like in [Mar+21]. Lastly, the current analysis also relied on a peak detection algorithm to identify the most probable timings of gait events [Gad+19; Mar+21; Rom+22]. However, from a clinical perspective, this may actually be regarded as a benefit, since it would allow a clinician to decide whether to take into account certain strides based on how confidently it can be assumed that it was indeed a stride.

5.5. 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 to extend the current analysis for data from multiple days, and to evaluate to which extent the DL network can be trained using participant-specific data to improve gait events detection on an individual level.

Acknowledgements

Foremost, the authors would like 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.

Funding

We acknowledge financial support by Land Schleswig-Holstein within the funding programme Open Access Publikationsfonds. This work 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 SDD 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.

LA, LR, AY and SDD were also supported by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust.

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.

This work 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".

Contributions

Conceptualization, R.R., F.S., C.H., A.C., G.S. and W.M.; methodology, R.R., F.S., A.C., G.S.; software, R.R., F.S.; validation, R.R.; formal analysis, R.R.; investigation, R.R.; resources, A.C., W.M.; data curation, R.R., F.S., A.C.; writing—original draft preparation, R.R.; writing—review and editing, F.S., C.H., A.C., G.S., W.M.; visualization, R.R.; supervision, G.S., W.M.; project administration, R.R., F.S., C.H., A.C., G.S. and W.M.; funding acquisition, A.C., W.M.; intellectual contribution: R.R., F.S., C.H., A.K., A.P., A.C., L.A., K.A., C.B., S.B., T.B., P.B., E.B., A.C., A.-E.C., M.C., B.C., L.C., I.D.'A. S.D.D., B.E., S.F., M.F., J.G.-A., E.G., J.H., H.H., E.H., A.K., C.K., F.K., S.K., C.M., D.M., E.M.-A., A.M., L.P., L.R., L.S., K.S., B.S., D.S., A.S., M.U., B.V., I.V., A.Y., G.S., W.M. All authors have read and agreed to the published version of the manuscript.

Ethics approval

Participants were recruited in five sites: Tel Aviv Sourasky Medical Center, Israel (ethics approval granted by the Helsinki Committee, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, 0551-19TLV), Robert Bosch Foundation for Medical Research, Germany (ethics approval granted by the ethical committee of the medical faculty of The University of T □ ubingen, 647/2019BO2), University of Kiel, Germany (ethics approval granted by the ethical committee of the medical faculty of Kiel University, D438/18), The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK and Sheffield Teaching Hospitals NHS Foundation Trust, UK (ethics approval granted by London – Bloomsbury Research Ethics committee, 19/LO/1507).

Part III

Discussion & Conclusion

6

General Discussion

6.1. Main Contributions

The aim of this thesis was to develop and validate an algorithm for robust and accurate gait event detection based on IMU signals, which holds in both supervised and unsupervised settings across healthy and different neurodegenerative disease cohorts. The focus was on IMU signals recorded from a shank-worn or foot-worn IMU as it is easier to detect gait events from these sensor positions [Sal+04; TCD14; Kon+16b; LYL11]. Therefore, an existing algorithm [Sal+04] was first implemented and validated for different walking tasks [Rom+21]. The algorithm has been described in research for adult [Sal+04] and elderly [Naj+09; Roc+10; See+10] as well as pathological gait (e.g., osteoarthritis [Ami+04] and Parkinson's disease [Sal+04; Sal+13]).

We have shown that the algorithm works well for a convenience sample comprised of older adults, people with Parkinson's disease and people who have had a stroke, but also found that there were more errors in gait events detection for steps during turns [Rom+21]. For home-based gait monitoring this may be a problem, as many daily life gait bouts are not purely straight-line walking and approximately $30\,\%$ is spent along curved trajectories [Gla+07; Tur+18]. Furthermore, the algorithm was based on identifying peaks in the mediolateral angular velocity signal. This requires knowing which sensor axis aligns with the mediolateral anatomical axis. In supervised settings this can easily be controlled for, but in remote monitoring participants often have to (re-)attach the sensor themselves, for example after showering. Therefore, our subsequent works have focused on methods that are invariant to sensor alignment, namely a data-driven deep learning approach based on convolutional neural networks [Rom+22].

Using a CNN-based deep neural network we have achieved a gait events detection performance that was on par or better than the heuristics-based algorithm [Rom+22]. The detection performance was comparable across four different sensor positions (i.e., left or right shank, and left or right ankle) and three different walking speeds for a heterogeneous study population with different neurological gait patterns, thereby showing its potential for use in remote gait monitoring. Subsequently, we used the trained model on a new dataset that was recorded in the habitual environment (e.g., home, work or community) and tested the performance against reference labels obtained from IMU and pressure insoles data. Like in the lab-based study, we achieved a high gait events detection rate across different mobility-limiting disease cohorts confirming that the deep learning algorithm is suitable for gait events detection even for heterogeneous gait patterns.

6.2. Discussion

6.2.1. Sensor Configuration

In the current thesis, a deep learning-based algorithm was developed and validated to detect gait events from IMU signals recorded with a shank-, ankle- or foot-worn device. Generally, these po-

sitions are better suited for detecting gait events than a low back-worn device [Pan+18; SBM16]. However, the low back is the clinically most favourable location for a single device [Hub+15; Mic+17; Del+16a], given

- its cost (one device),
- its location near to the centre of mass (which represents the overall human motion pattern),
- ergonomic conditions when worn attached to a belt or affixed to the skin,
- and its clinical value for fall risk, trunk stability and balance control.

In conventional approaches using a low back-worn device, gait events are detected either from the acceleration in the vertical direction [Naj+03; Hic+16; DGR16], anteroposterior direction [ZH03; Pha+17], or acceleration norm [Mas+15; Par+19], and are often based on identifying peaks in the corresponding acceleration signals. Therefore, algorithms based on a shank-worn device were not transferable to a low back-worn device. The current deep learning algorithm, however, can potentially be used to identify gait events from a low back-worn device without having to change the algorithm design given that the training dataset can be supplied with appropriately labelled data.

6.2.2. Sensor Alignment

Most conventional approaches heavily rely on accurate sensor-to-segment alignment. Inaccurate alignment estimation results in errors in following steps of gait analysis as most gait cycle parameters depend on gait events detection [Rue+10; Car+22]. There are four major approaches for dealing with sensor alignment [Pac+20; LGA19]:

- manual alignment,
- functional alignment,
- model-based alignment,
- and augmented alignment.

Manual alignment is based on attaching the IMU to the body segment such that the device visually aligns with the underlying anatomical axis. Functional alignment requires the study participant to hold a certain pose (to estimate the axis that aligns with gravity) or perform a predefined movement (to estimate the axis about which some joint rotates). Likewise, model-based alignment assume a model for the joint, e.g., a pin joint for the knee. Lastly, augmented alignment uses information from additional sources, such as an optical motion capture system. Clearly, the latter method is almost infeasible for remote gait monitoring, whereas manual alignment cannot be always controlled for in the habitual environment. Functional alignment mostly consists of holding a predefined pose and performing movements [Kon+16b] after which principal component analysis [Shl14] can be used to

estimate the sensitive axis that aligns with the mediolateral axis [Kon+16b; Mar+13; Car+22], and likewise assumes constraints on joint motion to estimate the mediolateral axis [SRS14]. Although these methods are valid, not all study participants are capable of holding these poses and performing the tasks [Fav+08; Fav+09; Cut+09], and the proposed deep learning algorithm requires no prior knowledge on the sensor alignment.

6.2.3. Generalizability

Given a deep learning model, one of the main concerns is how it generalizes to data outside of the dataset distribution. That said, the same concerns obviously apply to any heuristics-based decision algorithm. Within the field of deep learning, however, this concern has been addressed specifically. To accommodate data from outside the training data, so-called data augmentation techniques could be applied so that the algorithm experiences more variable data during training.

6.2.4. Explainability

While the current deep learning algorithm yielded very good results regarding gait event detection, it suffers from a limitation common to deep learning approaches, namely its black-box character [AB18; Hor+19]. Although the underlying mathematical principles are understood, it often remains unclear why a particular prediction has been made and if meaningful patterns have resulted in this prediction. This currently hinders the widespread adoption of DL-based decision support systems in medical practice [Hol+17]. Therefore, the field of explainable artificial intelligence (XAI) has gain increasing attention in recent years. XAI methods intend to illustrate how complex and non-linear ML models operate and how they produced their predictions [Sli+21]. A first step in XAI for clinical gait analysis explained predictions in gait-based person recognition [Hor+19] using layer-wise relevance propagation (LRP) [Bac+15; Mon+19]. Later, the authors used LRP to classify pathological gait from ground reaction forces [Sli+21]. Alternative approaches have used gait features derived from IMU data [Lor+12; GLR13] and then used recursive feature elimination (RFE) [Gra+06] to determine which gait features were most useful in classifying parkinsonian gait [Reh+19].

6.2.5. Context Awareness

It was suggested that monitoring gait in the *real world* with body-worn sensors, such as inertial measurement units, provides more sensitive clinical endpoints for diagnostic and therapeutic purposes [Del+16a; Gol+18]. The idea was that disease-specific signatures of gait impairments are stressed more in the real life due to the complexity and challenges associated with these environments [Ore+08].

However, differences in (spatio-)temporal gait parameters derived from inertial measurement units can only be interpreted with caution, as it will likely be infeasible to associate changes in gait

parameters with an underlying physiological changes and not with environmental changes, such as the surface terrain [Pat+14; Noh+21]. For example, it was shown that healthy adults have more walking activity *not at home* than at home, whereas this difference was negligible for Parkinson's disease patients [Rot+21b]. Furthermore, subjects from both cohorts tended to walk slower at home.

To capture the environment in which walking takes place, it may be an option to include additional sensing modalities. For example, a bluetooth gateway placed at home could register if inertial measurement units are within the local neighborhood, and thus walking takes place at home [Rot+21b], or data from a global positioning system could be used to extract repeated walking paths, so that only walking parameters from the same path are compared on an intra-individual basis [WA19].

6.3. Future Outlook

Currently, the developed algorithm was tested for shank-, ankle- and foot-worn IMUs across different neurological and healthy cohorts. Future studies may include other device locations, for example the low back. The low back is among the most clinically favourable location for a single device, given its cost (one device), its location near to the centre of mass (which represents the overall human motion pattern), ergonomic conditions when worn attached to a belt or affixed to the skin, and its clinical value for fall risk, trunk stability and balance control [Mic+17; Hub+15; Del+17].

Furthermore, the focus has been on identifying gait events to eventually gain insights into the habitual gait pattern. However, other activities of daily life (ADLs) are likely affected by disease as well, and may be an interesting clinical endpoint as well. For example, a wrist-worn IMU was used to quantify eating difficulties in PD [Kyr+21].

6.4. Conclusion

In conclusion, the aim was to develop and validate an algorithm for robust and accurate gait event detection based on IMU signals, that holds in both supervised and unsupervised settings across both healthy and different neurodegenerative disease cohorts. Therefore, a deep learning algorithm was developed and deployed both in supervised laboratory settings and in free-living settings. It was tested across heterogeneous cohorts of different neurological diseases and a healthy cohort. The algorithm showed excellent performance in gait events detection, and subsequently extracted stride-specific gait parameters matched with the values obtained with respective reference systems. Therefore, the algorithm provides an excellent modular framework to extract clinically relevant parameters related to the gait pattern obtained with inertial measurement units.

References

[AB18] A. Adadi and M. Berrada. “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)”. In: *IEEE Access* 6 (2018), pp. 52138–52160. DOI: 10.1109/ACCESS.2018.2870052.

[AB83] D.G. Altman and J.M. Bland. “Measurement in Medicine: The Analysis of Method Comparison Studies”. In: *The Statistician* 32.3 (1983), p. 307. DOI: 10.2307/2987937.

[Ada+17] J.L. Adams et al. “Multiple Wearable Sensors in Parkinson and Huntington Disease Individuals: A Pilot Study in Clinic and at Home”. In: *Digital Biomarkers* 1.1 (2017), pp. 52–63. DOI: 10.1159/000479018.

[All+19] E.K. Allseits et al. “Characterizing the Impact of Sampling Rate and Filter Design on the Morphology of Lower Limb Angular Velocities”. In: *IEEE Sensors Journal* 19.11 (2019), pp. 4115–4122.

[AM85] E.K. Antonsson and R.W. Mann. “The Frequency Content of Gait”. In: *Journal of Biomechanics* 18.1 (1985), pp. 39–47. DOI: 10.1016/0021-9290(85)90043-0.

[Ami+02] K. Aminian et al. “Spatio-Temporal Parameters of Gait Measured by an Ambulatory System Using Miniature Gyroscopes”. In: *Journal of Biomechanics* 35.5 (2002), pp. 689–699. DOI: 10.1016/S0021-9290(02)00008-8.

[Ami+04] K. Aminian et al. “Evaluation of an Ambulatory System for Gait Analysis in Hip Osteoarthritis and After Total Hip Replacement”. In: *Gait & Posture* 20.1 (2004), pp. 102–107. DOI: 10.1016/S0966-6362(03)00093-6.

[Ami+95] K. Aminian et al. “Incline, Speed, and Distance Assessment During Unconstrained Walking”. In: *Medicine & Science in Sports & Exercise* 27.2 (1995), pp. 226–234.

[Ank+15] M.M. Ankaralı et al. “Walking Dynamics Are Symmetric (Enough)”. In: *Journal of The Royal Society Interface* 12.20150209 (2015), p. 20150209. DOI: 10.1098/rsif.2015.0209.

[Atr+21] A. Atrsaei et al. “Gait Speed in Clinical and Daily Living Assessments in Parkinson’s Disease Patients: Performance Versus Capacity”. In: *NPJ Parkinson’s disease* 7.1 (2021), p. 24. DOI: 10.1038/s41531-021-00171-0.

[Bac+15] S. Bach et al. “On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation”. In: *PLOS ONE* 10.7 (2015), e0130140. DOI: 10.1371/journal.pone.0130140.

[Bae+13] T. Baetens et al. “Gait Analysis With Cognitive-Motor Dual Tasks to Distinguish Fallers From Nonfallers Among Rehabilitating Stroke Patients”. In: *Archives of Physical Medicine and Rehabilitation* 94.4 (2013), pp. 680–686. DOI: https://doi.org/10.1016/j.apmr.2012.11.023.

[BB12] J. Bergstra and Y. Bengio. “Random Search for Hyper-Parameter Optimization”. In: *Journal of Machine Learning Research* 13 (2012), pp. 281–305.

[BBS14] A. Bulling, U. Blanke, and B. Schiele. “A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors”. In: *ACM Computing Surveys* 46.3 (2014), 33:1–33:33. DOI: 10.1145/2499621.

[BDC19] S. Bertuletti, U. Della Croce, and A. Cereatti. “A Wearable Solution for Accurate Step Detection Based on the Direct Measurement of the Inter-Foot Distance”. In: *Journal of Biomechanics* 84 (2019), pp. 274–277. DOI: 10.1016/j.jbiomech.2018.12.039.

[Ber+17] S. Bertuletti et al. “Static and Dynamic Accuracy of an Innovative Miniaturized Wearable Platform for Short Range Distance Measurements for Human Movement Applications”. In: *Sensors* 17.7 (2017), p. 1492. DOI: 10.3390/s17071492.

[Ber+18] M. Bertoli 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”. In: *BioMedical Engineering OnLine* 17.1 (2018), p. 58. DOI: 10.1186/s12938-018-0488-2.

[Bis06] C.M. Bishop. *Pattern Recognition and Machine Learning*. Singapore: Springer Science+Business Media, LLC, 2006. ISBN: 978-0-387-31073-2.

[BKK18] S. Bai, J.Z. Kolter, and V. Koltun. *An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling*. 2018. DOI: 10.48550/ARXIV.1803.01271.

[Blo+00] B.R. Bloem et al. “Idiopathic Senile Gait Disorders Are Signs of Subclinical Disease”. In: *Journal of the American Geriatrics Society* 48.9 (2000), pp. 1098–1101. DOI: 10.1111/j.1532-5415.2000.tb04786.x.

[Blo+92] B.R. Bloem et al. “Investigation of Gait in Elderly Subjects over 88 Years of Age”. In: *Journal of Geriatric Psychiatry and Neurology* 5.2 (1992), pp. 78–84. DOI: 10.1177/002383099200500204.

[BMM18] E. Buckley, C. Mazzà, and A. McNeill. “A Systematic Review of the Gait Characteristics Associated with Cerebellar Ataxia”. In: *Gait & Posture* 60 (2018), pp. 154–163. DOI: 10.1016/j.gaitpost.2017.11.024.

[Bob+18] V.N. Bobić et al. “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, pp. 2284–2287. DOI: 10.1109/EMBC.2018.8512836.

[Bon+20] T. Bonci et al. “An Objective Methodology for the Selection of a Device for Continuous Mobility Assessment”. In: *Sensors* 20.22 (2020). DOI: 10.3390/s20226509.

[Bon+22] T. Bonci et al. “An Algorithm for Accurate Marker-Based Gait Event Detection in Healthy and Pathological Populations During Complex Motor Tasks”. In: *Frontiers in Bioengineering and Biotechnology* 10 (2022). DOI: 10.3389/fbioe.2022.868928.

[Böt+16] K. Bötzel et al. “Gait Recording with Inertial Sensors – How to Determine Initial and Terminal Contact”. In: *Journal of Biomechanics* 49.3 (2016), pp. 332–337. DOI: 10.1016/j.jbiomech.2015.12.035.

[BR14] D.A. Bruening and S.T. Ridge. “Automated Event Detection Algorithms in Pathological Gait”. In: *Gait & Posture* 39.1 (2014), pp. 472–477. DOI: 10.1016/j.gaitpost.2013.08.023.

[BRG15] K. Ben Mansour, N. Rezzoug, and P. Gorce. “Analysis of Several Methods and Inertial Sensors Locations to Assess Gait Parameters in Able-Bodied Subjects”. In: *Gait & Posture* 42.4 (2015), pp. 409–414. DOI: 10.1016/j.gaitpost.2015.05.020.

[Bru+09] S.M. Bruijn et al. “Statistical Precision and Sensitivity of Measures of Dynamic Gait Stability”. In: *Journal of Neuroscience Methods* 178.2 (2009), pp. 327–333. DOI: 10.1016/j.jneumeth.2008.12.015.

[Buc+19] C. Buckley et al. “The Role of Movement Analysis in Diagnosing and Monitoring Neurodegenerative Conditions: Insights from Gait and Postural Control”. In: *Brain Sciences* 9.2 (2019). DOI: 10.3390/brainsci9020034.

[But+82] R.J. Butland et al. “Two-, Six-, and 12-Minute Walking Tests in Respiratory Disease”. In: *British Medical Journal (Clinical Research Ed.)* 284.6329 (1982), pp. 1607–1608. DOI: 10.1136/bmj.284.6329.1607.

[Cai+17] S.M. Cain et al. “Accurate and Robust Gait Event Detection Using Foot-Mounted Inertial Measurement Units”. In: *41st Annual Meeting of the American Society of Biomechanics*. 2017.

[Cam+18] J. Camps et al. “Deep Learning for Freezing of Gait Detection in Parkinson’s Disease Patients in Their Homes Using a Waist-Worn Inertial Measurement Unit”. In: *Knowledge-Based Systems* 139 (2018), pp. 119–131. DOI: 10.1016/j.knosys.2017.10.017.

[Car+18] L. Carcreff et al. “What is the Best Configuration of Wearable Sensors to Measure Spatiotemporal Gait Parameters in Children with Cerebral Palsy?” In: *Sensors* 18.2 (2018). DOI: 10.3390/s18020394.

[Car+19] L. Carcreff et al. “A Personalized Approach to Improve Walking Detection in Real-Life Settings: Application to Children with Cerebral Palsy”. In: *Sensors* 19.23 (2019). DOI: 10.3390/s19235316.

[Car+22] L. Carcreff et al. “Simple Rule to Automatically Recognize the Orientation of the Sagittal Plane Foot Angular Velocity for Gait Analysis Using IMUs on the Feet of Individuals with Heterogeneous Motor Disabilities”. In: *Journal of Biomechanics* 135 (2022), p. 111055. DOI: https://doi.org/10.1016/j.jbiomech.2022.111055.

[Car20] L. Carcreff. “Gait Analysis in Children with Cerebral Palsy: Bridging the Gap Between the Laboratory and Real Life”. PhD thesis. École polytechnique fédérale de Lausanne, 2020.

[Cat+08] P. Catalfamo et al. “Detection of Gait Events Using an F-Scan In-Shoe Pressure Measurement System”. In: *Gait & Posture* 28.3 (2008), pp. 420–426. DOI: 10.1016/j.gaitpost.2008.01.019.

[Cel+21] Y. Celik et al. “Gait Analysis in Neurological Populations: Progression in the Use of Wearables”. In: *Medical Engineering & Physics* 87 (2021), pp. 9–29. DOI: 10.1016/j.medengphy.2020.11.005.

[Cer18] F. Cerreta. *EMA Experience on mHealth Technology*. https://www.ema.europa.eu/en/documents/presentation/presentation-ema-experience-mhealth-technology-emas-pcwp-hcpwp-joint-meeting-francesca-cerreta_en.pdf. 2018.

[CGE10] P. Catalfamo, S. Ghoussayni, and D. Ewins. “Gait Event Detection on Level Ground and Incline Walking Using a Rate Gyroscope”. In: *Sensors* 10.6 (2010), pp. 5683–5702. DOI: 10.3390/s100605683.

[Chi+05] L. Chiari et al. “Human Movement Analysis Using Stereophotogrammetry: Part 2: Instrumental Errors”. In: *Gait & Posture* 21.2 (2005), pp. 197–211. DOI: 10.1016/j.gaitpost.2004.04.004.

[Cho+15] F. Chollet et al. *Keras*. https://github.com/keras-team/keras. 2015.

[Cho21] F. Chollet. *Deep Learning with Python*. 2nd. Shelter Island, NY, USA: Manning Publications, 2021.

[Col+17] L. Colin et al. “Temporal Convolutional Networks for Action Segmentation and Detection”. In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. Institute of Electrical and Electronics Engineers Inc., 2017, pp. 1003–1012. DOI: 10.1109/CVPR.2017.113.

[Cor+19] P. Coran et al. “Advancing the Use of Mobile Technologies in Clinical Trials: Recommendations from the Clinical Trials Transformation Initiative”. In: *Digital Biomarkers* 3.3 (2019), pp. 145–154. DOI: 10.1159/000503957.

[Cor+21] M.F. Corrà et al. “Comparison of Laboratory and Daily-Life Gait Speed Assessment during ON and OFF States in Parkinson’s Disease”. In: *Sensors* 21 (2021). DOI: 10.3390/s21123974.

[CS02] H.G. Chambers and D.H. Sutherland. “A Practical Guide to Gait Analysis”. In: *Journal of the American Academy of Orthopaedic Surgeons* 10.3 (2002), pp. 222–231.

[Cuo+07] A.V. Cuomo et al. “Health-Related Quality of Life Outcomes Improve After Multilevel Surgery in Ambulatory Children With Cerebral Palsy”. In: *Journal of Pediatric Orthopaedics* 27.6 (2007), pp. 653–657. DOI: 10.1097/BPO.0b013e3180dca147.

[Cut+09] A.G. Cutti et al. “‘Outwalk’: a Protocol for Clinical Gait Analysis Based on Inertial and Magnetic Sensors”. In: *Medical & Biological Engineering & Computing* 48.1 (2009), p. 17. DOI: 10.1007/s11517-009-0545-x.

[DÇB19] D. Diez, M. Çetinkaya-Rundel, and C.D. Barr. *OpenIntro Statistics*. 4th. OpenIntro, Inc., 2019.

[DCM18] U. Della Croce, A. Cereatti, and M. Mancini. “Handbook of Human Motion”. In: Springer, Cham, 2018. Chap. Gait Parameters Estimated Using Inertial Measurement Units.

[de +10] W.H.K. de Vries et al. “Functionally Interpretable Local Coordinate Systems for the Upper Extremity Using Inertial & Magnetic Measurement Systems”. In: *Journal of Biomechanics* 43.10 (2010), pp. 1983–1988. DOI: 10.1016/j.jbiomech.2010.03.007.

[Del+16a] S. Del Din et al. “Free□Living Monitoring of Parkinson’s Disease: Lessons from the Field”. In: *Movement Disorders* 31.9 (2016), pp. 1293–1313. DOI: 10.1002/mds.26718.

[Del+16b] S. Del Din et al. “Instrumented Gait Assessment with a Single Wearable: An Introductory Tutorial”. In: *F1000Research* 5 (2016), p. 2323. DOI: 10.12688/f1000research.9591.1.

[Del+17] S. Del Din et al. “Free-Living Gait Characteristics in Ageing and Parkinson’s Disease: Impact of Environment and Ambulatory Bout Length”. In: *Journal of NeuroEngineering and Rehabilitation* 13.1 (2017), p. 46. DOI: 10.1186/s12984-016-0154-5.

[Del+19] S. Del Din et al. “Gait Analysis with Wearables Predicts Conversion to Parkinson Disease”. In: *Annals of Neurology* 86.3 (2019), pp. 357–367. DOI: 10.1002/ana.25548.

[DGR16] S. Del Din, A. Godfrey, and L. Rochester. “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”. In: *IEEE Journal of Biomedical and Health Informatics* 20.3 (2016), pp. 838–847. DOI: 10.1109/JBHI.2015.2419317.

[Dic10] J. Dicharry. “Kinematics and Kinetics of Gait: From Lab to Clinic”. In: *Clinics in Sports Medicine*. The Runner 29.3 (2010), pp. 347–364. DOI: 10.1016/j.csm.2010.03.013.

[DKZ10] B. Dijkstra, Y.P. Kamsma, and W. Zijlstra. “Detection of Gait and Postures Using a Miniaturised Triaxial Accelerometer-Based System: Accuracy in Community-Dwelling Older Adults”. In: *Age and Ageing* 39.2 (2010), pp. 259–262. DOI: 10.1093/ageing/afp249.

[Esk+16] B.M. Eskofier 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, pp. 655–658. DOI: 10.1109/EMBC.2016.7590787.

[Esp+19] A.J. Espay et al. “A Roadmap for Implementation of Patient-Centered Digital Outcome Measures in Parkinson’s Disease Obtained Using Mobile Health Technologies”. In: *Movement Disorders* 34.5 (2019), pp. 657–663. DOI: 10.1002/mds.27671.

[Fav+08] J. Favre et al. “Ambulatory Measurement of 3D Knee Joint Angle”. In: *Journal of Biomechanics* 41.5 (2008), pp. 1029–1035. DOI: 10.1016/j.jbiomech.2007.12.003.

[Fav+09] J. Favre et al. “Functional Calibration Procedure for 3D Knee Joint Angle Description Using Inertial Sensors”. In: *Journal of Biomechanics* 42.14 (2009), pp. 2330–2335. DOI: 10.1016/j.jbiomech.2009.06.025.

[Fed13] P.A. Federolf. “A Novel Approach to Solve the “Missing Marker Problem” in Marker-Based Motion Analysis That Exploits the Segment Coordination Patterns in Multi-Limb Motion Data”. In: *PLOS ONE* 8.10 (2013), pp. 1–13. DOI: 10.1371/journal.pone.0078689.

[FGP95] F. Ferraris, U. Grimaldi, and M. Parvis. “Procedure for Effortless In-Field Calibration of Three-Axis Rate Gyros and Accelerometers”. In: *Sensor and Materials* 7.5 (1995), pp. 311–330.

[Fig+10] D. Figo et al. “Preprocessing Techniques for Context Recognition from Accelerometer Data”. In: *Personal and Ubiquitous Computing* 14.7 (2010), pp. 645–662. DOI: 10.1007/s00779-010-0293-9.

[Fil+20] B. Filtjens et al. “A Data-Driven Approach for Detecting Gait Events During Turning in People with Parkinson’s Disease and Freezing of Gait”. In: *Gait & Posture* 80 (2020), pp. 130–136. DOI: 10.1016/j.gaitpost.2020.05.026.

[Fis+16] J.M. Fisher et al. “Body-Worn Sensors in Parkinson’s Disease: Evaluating Their Acceptability to Patients”. In: *Telemedicine and e-Health* 22.1 (2016), pp. 63–69. DOI: 10.1089/tmj.2015.0026.

[Fla+19] F. Flachenecker et al. “Objective Sensor-Based Gait Measures Reflect Motor Impairment in Multiple Sclerosis Patients: Reliability and Clinical Validation of a Wearable Sensor Device”. In: *Multiple Sclerosis and Related Disorders* 39 (2019), p. 101903. DOI: 10.1016/j.msard.2019.101903.

[FM20] A. Fasano and M. Mancini. “Wearable-Based Mobility Monitoring: The Long Road Ahead”. In: *The Lancet Neurology* 19.5 (2020), pp. 378–379. DOI: 10.1016/S1474-4422(20)30033-8.

[Fuk80] K. Fukushima. “Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position”. In: *Biological Cybernetics* 36.4 (1980), pp. 193–202. DOI: 10.1007/BF00344251.

[Gad+19] M. Gadaleta et al. “Deep Learning Techniques for Improving Digital Gait Segmentation”. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*. 2019, pp. 1834–1837. DOI: 10.1109/EMBC.2019.8856685.

[Gau+18] E.B. Gausden et al. “Gait Analysis After Intertrochanteric Hip Fracture: Does Shortening Result in Gait Impairment?” In: *Journal of Orthopaedic Trauma* 32.11 (2018), pp. 554–558. DOI: 10.1097/BOT.0000000000001283.

[GBC16] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[Gér19] A. Géron. *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*. 2nd. O’Reilly Media, Inc., 2019.

[GF16] Ø. Gløersen and P. Federolf. “Predicting Missing Marker Trajectories in Human Motion Data Using Marker Intercorrelations”. In: *PLOS ONE* 11.3 (2016), pp. 1–14. DOI: 10.1371/journal.pone.0152616.

[Gia15] D. Giavarina. “Understanding Bland Altman Analysis”. In: *Biochemia Medica* 25.2 (2015), pp. 141–151. DOI: 10.11613/BM.2015.015.

[GL88] W.R. Gibb and A.J. Lees. “The Relevance of the Lewy Body to the Pathogenesis of Idiopathic Parkinson’s Disease”. In: *Journal of Neurology, Neurosurgery, and Psychiatry* 51.6 (1988), pp. 745–752. DOI: 10.1136/jnnp.51.6.745.

[Gla+07] B.C. Glaister et al. “Video Task Analysis of Turning During Activities of Daily Living”. In: *Gait & Posture* 25.2 (2007), pp. 289–294. DOI: 10.1016/j.gaitpost.2006.04.003.

[GLR13] B. Galna, S. Lord, and L. Rochester. “Is Gait Variability Reliable in Older Adults and Parkinson’s Disease? Towards an Optimal Testing Protocol”. In: *Gait & Posture* 37.4 (2013), pp. 580–585. DOI: https://doi.org/10.1016/j.gaitpost.2012.09.025.

[God+08] A. Godfrey et al. “Direct Measurement of Human Movement by Accelerometry”. In: *Medical Engineering & Physics* 30.10 (2008), pp. 1364–1386. DOI: 10.1016/j.medengphy.2008.09.005.

[God+16] C. Godinho et al. “A Systematic Review of the Characteristics and Validity of Monitoring Technologies to Assess Parkinson’s Disease”. In: *Journal of NeuroEngineering and Rehabilitation* 13.1 (2016), p. 24. DOI: 10.1186/s12984-016-0136-7.

[God+22] F.E. Godkin et al. “Feasibility of a Continuous, Multi-Sensor Remote Health Monitoring Approach in Persons Living with Neurodegenerative Disease”. In: *Journal of Neurology* 269.5 (2022), pp. 2673–2686. DOI: 10.1007/s00415-021-10831-z.

[Goe+08] C.G. Goetz et al. “Movement Disorder Society-Sponsored Revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS): Scale Presentation and Clinimetric Testing Results”. In: *Movement Disorders* 23.15 (2008), pp. 2129–2170. DOI: 10.1002/mds.22340.

[Gol+18] M. Gold et al. “Digital Technologies as Biomarkers, Clinical Outcomes Assessment, and Recruitment Tools in Alzheimer’s Disease Clinical Trials”. In: *Alzheimer’s & Dementia: Translational Research & Clinical Interventions* 4.1 (2018), pp. 234–242. DOI: 10.1016/j.trci.2018.04.003.

[Gra+06] P.M. Granitto et al. “Recursive Feature Elimination with Random Forest for PTR-MS Analysis of Agroindustrial Products”. In: *Chemometrics and Intelligent Laboratory Systems* 83.2 (2006), pp. 83–90. DOI: 10.1016/j.chemolab.2006.01.007.

[Gre+10] B.R. Greene et al. “An Adaptive Gyroscope-Based Algorithm for Temporal Gait Analysis”. In: *Medical & Biological Engineering & Computing* 48 (2010), pp. 1251–1260. DOI: 10.1007/s11517-010-0692-0.

[Gri+16] J. Griškevičius et al. “Estimation of Temporal Gait Parameters of Multiple Sclerosis Patients in Clinical Setting Using Inertial Sensors”. In: *Proceedings of the 11th International Conference BIOMDLORE 2016*. 2016, pp. 80–82.

[Gur+94] J.M. Guralnik et al. “A Short Physical Performance Battery Assessing Lower Extremity Function: Association with Self-Reported Disability and Prediction of Mortality and Nursing Home Admission”. In: *Journal of Gerontology* 49.2 (1994), pp. M85–94. DOI: 10.1093/geronj/49.2.m85.

[HA05] J.M. Hausdorff and N.B. Alexander. *Gait Disorders: Evaluation and Management*. 1st. Boca Raton, FL, USA: CRC Press, 2005.

[Haj+18] N. Haji Ghassemi et al. “Segmentation of Gait Sequences in Sensor-Based Movement Analysis: A Comparison of Methods in Parkinson’s Disease”. In: *Sensors* 18.1 (2018). DOI: 10.3390/s18010145.

[Han+16] J. Hannink et al. “Sensor-Based Gait Parameter Extraction With Deep Convolutional Neural Networks”. In: *IEEE Journal of Biomedical and Health Informatics* 21.1 (2016), pp. 85–93. DOI: 10.1109/JBHI.2016.2636456.

[Hau05] J.M. Hausdorff. “Gait Variability: Methods, Modeling and Meaning”. In: *Journal of NeuroEngineering and Rehabilitation* 2.1 (2005), p. 19. DOI: 10.1186/1743-0003-2-19.

[He+16] K. He et al. “Deep Residual Learning for Image Recognition”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[HHP16] Y. Hammerla N, S. Halloran, and T. Plötz. “Deep, Convolutional, and Recurrent Models for Human Activity Recognition Using Wearables | Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence”. In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*. New York, New York, USA, 2016, pp. 1533–1540. DOI: 10.5555/3060832.3060835.

[Hic+16] A. Hickey et al. “Detecting Free-Living Steps and Walking Bouts: Validating an Algorithm for Macro Gait Analysis”. In: *Physiological Measurement* 38 (2016), N1. DOI: 10.1088/1361-6579/38/1/N1.

[Hil+19] I. Hillel 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”. In: *European Review of Aging and Physical Activity* 16.1 (2019), p. 6. DOI: 10.1186/s11556-019-0214-5.

[HM00] A. Hreljac and R.N. Marschall. “Algorithms to Determine Event Timing During Normal Walking Using Kinematic Data”. In: *Journal of Biomechanics* 33.6 (2000), pp. 783–786. DOI: 10.1016/S0021-9290(00)00014-2.

[Hod08] D. Hodgins. “The Importance of Measuring Human Gait”. In: *Medical device technology* 19.5 (2008), pp. 44–47.

[Hol+17] A. Holzinger et al. *What Do We Need to Build Explainable AI Systems for the Medical Domain?* 2017. DOI: 10.48550/arXiv.1712.09923.

[Hor+19] F. Horst et al. “Explaining the Unique Nature of Individual Gait Patterns With Deep Learning”. In: *Scientific Reports* 9.1 (2019), p. 2391. DOI: 10.1038/s41598-019-38748-8.

[HS99] K. Hase and R.B. Stein. “Turning Strategies During Human Walking”. In: *Journal of Neurophysiology* 81.6 (1999), pp. 2914–2922. DOI: 10.1152/jn.1999.81.6.2914.

[HT82] A. Henik and J. Tzelgov. “Is Three Greater Than Five: The Relation Between Physical and Semantic Size in Comparison Tasks”. In: *Memory & Cognition* 10.4 (1982), pp. 389–395. DOI: 10.3758/BF03202431.

[Hub+15] R.P. Hubble et al. “Wearable Sensor Use for Assessing Standing Balance and Walking Stability in People with Parkinson’s Disease: A Systematic Review”. In: *PLOS ONE* 10.4 (2015), pp. 1–22. DOI: 10.1371/journal.pone.0123705.

[Hui+18] B. Huijben et al. “The Effect of Walking Speed on Quality of Gait in Older Adults”. In: *Gait & Posture* 65 (2018), pp. 112–116. DOI: 10.1016/j.gaitpost.2018.07.004.

[Ilu+14] T. Iluz 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”. In: *Journal of NeuroEngineering and Rehabilitation* 11.1 (2014), p. 48. DOI: 10.1186/1743-0003-11-48.

[Ios+16] M. Iosa et al. “Wearable Inertial Sensors for Human Movement Analysis”. In: *Expert Review of Medical Devices* 13.7 (2016), pp. 641–659. DOI: 10.1080/17434440.2016.1198694.

[IS15] S. Ioffe and C. Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. 2015. DOI: 10.48550/ARXIV.1502.03167. URL: https://arxiv.org/abs/1502.03167.

[Jar+18] D. Jarchi et al. “A Review on Accelerometry-Based Gait Analysis and Emerging Clinical Applications”. In: *IEEE Reviews in Biomedical Engineering* 11 (2018), pp. 177–194. DOI: 10.1109/RBME.2018.2807182.

[Jas+06] J.M. Jasiewicz et al. “Gait Event Detection Using Linear Accelerometers or Angular Velocity Transducers in Able-Bodied and Spinal-Cord Injured Individuals”. In: *Gait & Posture* 24.4 (2006), pp. 502–509. DOI: 10.1016/j.gaitpost.2005.12.017.

[Ji+19] N. Ji et al. “Appropriate Mother Wavelets for Continuous Gait Event Detection Based on Time-Frequency Analysis for Hemiplegic and Healthy Individuals”. In: *Sensors* 19.16 (2019). DOI: 10.3390/s19163462.

[Kar+06] D.M. Karantonis et al. “Implementation of a Real-Time Human Movement Classifier Using a Triaxial Accelerometer for Ambulatory Monitoring”. In: *IEEE Transactions on Information Technology in Biomedicine* 10.1 (2006), pp. 156–167. DOI: 10.1109/TITB.2005.856864.

[KB14] D.P. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. 2014. DOI: 10.48550/ARXIV.1412.6980. URL: https://arxiv.org/abs/1412.6980.

[KDS19] Ł. Kidziński, S. Delp, and M. Schwartz. “Automatic Real-Time Gait Event Detection in Children Using Deep Neural Networks”. In: *PLoS ONE* 14.1 (2019), e0211466. DOI: 10.1371/journal.pone.0211466.

[KHS17] M. Kok, J.D. Hol, and J.D. Schön. “Using Inertial Sensors for Position and Orientation Estimation”. In: *Foundations and Trends® in Signal Processing* 11.1–2 (2017), pp. 1–153. DOI: 10.1561/2000000094.

[KJ74] J. Kormylo and V. Jain. “Two-Pass Recursive Digital Filter with Zero Phase Shift”. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 22.5 (1974), pp. 384–387. DOI: 10.1109/TASSP.1974.1162602.

[Klu+21] F. Kluge et al. “Consensus Based Framework for Digital Mobility Monitoring”. In: *PLOS One* 16.8 (2021), e0256541. DOI: 10.1371/journal.pone.0256541.

[Kon+16a] W. Kong et al. “Anatomical Calibration through Post-Processing of Standard Motion Tests Data”. In: *Sensors* 16.12 (2016), p. 2011. DOI: 10.3390/s16122011.

[Kon+16b] W. Kong et al. “Comparison of Gait Event Detection from Shanks and Feet in Single-Task and Multi-Task Walking of Healthy Older Adults”. In: *2016 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. 2016, pp. 2063–2068. DOI: 10.1109/ROBIO.2016.7866633.

[Kön+17] A. König et al. “Objective Measurement of Gait Parameters in Healthy and Cognitively Impaired Elderly Using the Dual-Task Paradigm”. In: *Aging Clinical and Experimental Research* 29.6 (2017), pp. 1181–1189. DOI: 10.1007/s40520-016-0703-6.

[KW16] S. Khandelwal and N. Wickström. “Gait Event Detection in Real-World Environment for Long-Term Applications: Incorporating Domain Knowledge Into Time-Frequency Analysis”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 24.12 (2016), pp. 1363–1372. DOI: 10.1109/TNSRE.2016.2536278.

[Kyr+21] K. Kyritsis et al. “Assessment of real life eating difficulties in Parkinson’s disease patients by measuring plate to mouth movement elongation with inertial sensors”. In: *Scientific Reports* 11.1 (2021), p. 1632. DOI: 10.1038/s41598-020-80394-y.

[Lam+16] S. Lambrecht et al. “Inertial Sensor Error Reduction through Calibration and Sensor Fusion”. In: *Sensors (Basel, Switzerland)* 16.2 (2016), p. 235. DOI: 10.3390/s16020235.

[LBH15] Y. LeCun, Y. Bengio, and G. Hinton. “Deep Learning”. In: *Nature* 521.7553 (2015), pp. 436–444. DOI: 10.1038/nature14539.

[LBH95] Y. LeCun, Y. Bengio, and G. Hinton. “Deep Learning”. In: *Nature* 521.7553 (1995), pp. 436–444. DOI: 10.1038/nature14539.

[LBM02] C.J.C. Lamoth, P.J. Beek, and O.G. Meijer. “Pelvis–Thorax Coordination in the Transverse Plane During Gait”. In: *Gait & Posture* 16.2 (2002), pp. 101–114. DOI: 10.1016/S0966-6362(01)00146-1.

[LeC+98] Y. LeCun et al. “Gradient-Based Learning Applied to Document Recognition”. In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324. DOI: 10.1109/5.726791.

[Lem+20] M. Lempereur 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”. In: *Journal of Biomechanics* 98 (2020), p. 109490. DOI: 10.1016/j.jbiomech.2019.109490.

[LGA19] M.J. Leineweber, M.D. Gomez Orozco, and J. Andrysek. “Evaluating the Feasibility of Two Post-Hoc Correction Techniques for Mitigating Posture-Induced Measurement Errors Associated with Wearable Motion Capture”. In: *Medical Engineering & Physics* 71 (2019), pp. 38–44. DOI: https://doi.org/10.1016/j.medengphy.2019.06.013.

[Li+22] Y. Li et al. *What Makes Convolutional Models Great on Long Sequence Modeling?* 2022. DOI: 10.48550/ARXIV.2210.09298.

[Liu+19] T. Liu et al. “Towards Understanding the Importance of Shortcut Connections in Residual Networks”. In: *CoRR* abs/1909.04653 (2019).

[Lön+19] M. Löning et al. *sktime: A Unified Interface for Machine Learning with Time Series*. 2019. DOI: 10.48550/ARXIV.1909.07872.

[Lor+12] S. Lord et al. “Independent Domains of Gait in Older Adults and Associated Motor and Nonmotor Attributes: Validation of a Factor Analysis Approach”. In: *The Journals of Gerontology: Series A* 68.7 (2012), pp. 820–827. DOI: 10.1093/gerona/gls255.

[LRW12] D. Levine, J. Richards, and M. W. Whittle. “Whittle’s Gait Analysis”. In: 5th. London, UK: Churchill Livingstone, 2012. Chap. Normal Gait, pp. 29–64.

[LYL11] A. Laudanski, S. Yang, and Q. Li. “A Concurrent Comparison of Inertia Sensor-Based Walking Speed Estimation Methods”. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011, pp. 3484–3487. DOI: 10.1109/IEMBS.2011.6090941.

[Mae+21] W. Maetzler et al. “Modernizing Daily Function Assessment in Parkinson’s Disease Using Capacity, Perception, and Performance Measures”. In: *Movement Disorders* 36.1 (2021), pp. 76–82. DOI: 10.1002/mds.28377.

[Maq+16] H.F. Maqbool et al. “Real-Time Gait Event Detection for Lower Limb Amputees Using a Single Wearable Sensor”. In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2016, pp. 5067–5070. DOI: 10.1109/EMBC.2016.7591866.

[Mar+02] S. Marquis et al. “Independent Predictors of Cognitive Decline in Healthy Elderly Persons”. In: *Archives of Neurology* 59.4 (2002), pp. 601–606. DOI: 10.1001/archneur.59.4.601.

[Mar+13] B. Mariani et al. “Quantitative Estimation of Foot-Flat and Stance Phase of Gait Using Foot-Worn Inertial Sensors”. In: *Gait & Posture* 37.2 (2013), pp. 229–234. DOI: 10.1016/j.gaitpost.2012.07.012.

[Mar+17] C.F. Martindale 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)*. 2017, pp. 1266–1269. DOI: 10.1109/EMBC.2017.8037062.

[Mar+21] C.F. Martindale et al. “Wearables-Based Multi-Task Gait and Activity Segmentation Using Recurrent Neural Networks”. In: *Neurocomputing* 432 (2021), pp. 250–261. DOI: 10.1016/j.neucom.2020.08.079.

[Mas+15] F. Massé et al. “Improving Activity Recognition Using a Wearable Barometric Pressure Sensor in Mobility-Impaired Stroke Patients”. In: *Journal of NeuroEngineering and Rehabilitation* 12.1 (2015), p. 72. DOI: 10.1186/s12984-015-0060-2.

[Maz+21] C. Mazzà et al. “Technical Validation of Real-World Monitoring of Gait: A Multicentric Observational Study”. In: *BMJ Open* 11.12 (2021), e050785. DOI: 10.1136/bmjopen-2021-050785.

[McC+12] J. McCamley et al. “An Enhanced Estimate of Initial Contact and Final Contact Instants of Time Using Lower Trunk Inertial Sensor Data”. In: *Gait & Posture* 36.2 (2012), pp. 316–318. DOI: 10.1016/j.gaitpost.2012.02.019.

[McD+01] A.L. McDonough et al. “The Validity and Reliability of the GAITRite System’s Measurements: A Preliminary Evaluation”. In: *Archives of Physical Medicine and Rehabilitation* 82.3 (2001), pp. 419–425. DOI: 10.1053/apmr.2001.19778.

[Mei20] M. Meijs. “Automated Image Analysis and Machine Learning to Detect Cerebral Vascular Pathology in 4D-CTA”. PhD thesis. Radboud Universiteit Nijmegen, 2020.

[MFL15] A. Middleton, S.L. Fritz, and M. Lusardi. “Walking Speed: The Functional Vital Sign”. In: *Journal of Aging and Physical Activity* 23.2 (2015), pp. 314–322. DOI: 10.1123/japa.2013-0236.

[MGM14] 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”. In: *Sensors* 14.2 (2014), pp. 3362–3394. DOI: 10.3390/s140203362.

[Mic+17] E.M. Micó-Amigo et al. “Is the Assessment of 5 Meters of Gait with a Single Body-Fixed-Sensor Enough to Recognize Idiopathic Parkinson’s Disease-Associated Gait?” In: *Annals of Biomedical Engineering* 45 (2017), pp. 1266–1278. DOI: 10.1007/s10439-017-1794-8.

[Mil+15] A. Millecamps et al. “Understanding the Effects of Pre-Processing on Extracted Signal Features from Gait Accelerometry Signals”. In: *Computers in Biology and Medicine* 62 (2015), pp. 164–174. DOI: 10.1016/j.compbiomed.2015.03.027.

[Mil09] A. Miller. “Gait Event Detection Using a Multilayer Neural Network”. In: *Gait & Posture* 29.4 (2009), pp. 542–545. DOI: 10.1016/j.gaitpost.2008.12.003.

[Mir+19] A. Mirelman et al. “Gait Impairments in Parkinson’s Disease”. In: *The Lancet Neurology* 18.7 (2019), pp. 697–708. DOI: 10.1016/S1474-4422(19)30044-4.

[Moh+21] D.M. Mohan et al. “Assessment Methods of Post-stroke Gait: A Scoping Review of Technology-Driven Approaches to Gait Characterization and Analysis”. In: *Fronties in Neurology* 12 (2021), p. 650024. DOI: 10.3389/fneur.2021.650024.

[Mon+19] G. Montavon et al. “Layer-Wise Relevance Propagation: An Overview”. In: *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Ed. by W. Samek et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, pp. 193–209. DOI: 10.1007/978-3-030-28954-6_10.

[Mot+17] R.W. Motl et al. “Validity of the Timed 25-Foot Walk as an Ambulatory Performance Outcome Measure for Multiple Sclerosis”. In: *Multiple Sclerosis (Houndmills, Basingstoke, England)* 23.5 (2017), pp. 704–710. DOI: 10.1177/1352458517690823.

[MS11] A. Mannini and A.M. Sabatini. “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, pp. 4369–4373. DOI: 10.1109/IEMBS.2011.6091084.

[MS22] L. Moroney and E. Shyu. *Advanced Computer Vision with TensorFlow [MOOC]*. https://www.coursera.o computer-vision-with-tensorflow/home/welcome. 2022.

[Muñ+19] B. Muñoz Ospina et al. “Age Matters: Objective Gait Assessment in Early Parkinson’s Disease Using an RGB-D Camera”. In: *Parkinson’s Disease* 2019 (2019), e5050182. DOI: 10.1155/2019/5050182.

[Naj+03] B. Najafi et al. “Ambulatory System for Human Motion Analysis Using a Kinematic Sensor: Monitoring of Daily Physical Activity in the Elderly”. In: *IEEE transactions on bio-medical engineering* 50.6 (2003), pp. 711–723. DOI: 10.1109/TBME.2003.812189.

[Naj+09] B. Najafi et al. “Does Walking Strategy in Older People Change as a Function of Walking Distance?” In: *Gait & Posture* 29.2 (2009), pp. 261–266. DOI: 10.1016/j.gaitpost.2008.09.002.

[Nas+05] Z.S. Nasreddine et al. “The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment”. In: *Journal of the American Geriatrics Society* 53.4 (2005), pp. 695–699. DOI: https://doi.org/10.1111/j.1532-5415.2005.53221.x.

[Nie03] J.B. Nielsen. “How We Walk: Central Control of Muscle Activity during Human Walking”. In: *The Neuroscientist* 9.3 (2003), pp. 195–204. DOI: 10.1177/1073858403009003012.

[NK21] W. Niswander and K. Kontson. “Evaluating the Impact of IMU Sensor Location and Walking Task on Accuracy of Gait Event Detection Algorithms”. In: *Sensors* 21.12 (2021), p. 3989. DOI: 10.3390/s21123989.

[Noh+21] D. Nohelova et al. “Gait Variability and Complexity during Single and Dual-Task Walking on Different Surfaces in Outdoor Environment”. In: *Sensors* 21.14 (2021), p. 4792. DOI: 10.3390/s21144792.

[OCo+07] C.M. O’Connor et al. “Automatic Detection of Gait Events Using Kinematic Data”. In: *Gait & Posture* 25.3 (2007), pp. 469–474. DOI: 10.1016/j.gaitpost.2006.05.016.

[OMa+19] T. O’Malley et al. *KerasTuner*. https://github.com/keras-team/keras-tuner. 2019.

[Ore+08] M.S. Orendurff et al. “How Humans Walk: Bout Duration, Steps per Bout, and Rest Duration”. In: *Journal of Rehabilitation Research & Development* 45.7 (2008), pp. 1077–1089. DOI: 10.1682/JRRD.2007.11.0197.

[Pac+20] L. Pacher et al. “Sensor-to-Segment Calibration Methodologies for Lower-Body Kinematic Analysis with Inertial Sensors: A Systematic Review”. In: *Sensors* 20.11 (2020), p. 3322. DOI: 10.3390/s20113322.

[Pan+13] G. Panahandeh et al. “Continuous Hidden Markov Model for Pedestrian Activity Classification and Gait Analysis”. In: *IEEE Transactions on Instrumentation and Measurement* 62.5 (2013), pp. 1073–1083. DOI: 10.1109/TIM.2012.2236792.

[Pan+18] G.P. Panebianco et al. “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”. In: *Gait & Posture* 66 (2018), pp. 76–82. DOI: 10.1016/j.gaitpost.2018.08.025.

[Par+04] A. Paraschiv-Ionescu et al. “Ambulatory System for the Quantitative and Qualitative Analysis of Gait and Posture in Chronic Pain Patients Treated with Spinal Cord Stimulation”. In: *Gait & Posture* 20.2 (2004), pp. 113–125. DOI: 10.1016/j.gaitpost.2003.07.005.

[Par+19] A. Paraschiv-Ionescu et al. “Locomotion and Cadence Detection Using a Single Trunk-Fixed Accelerometer: Validity for Children with Cerebral Palsy in Daily Life-Like Conditions”. In: *Journal of Neuroengineering and Rehabilitation* 16.1 (2019), p. 24. DOI: 10.1186/s12984-019-0494-z.

[Pat+14] M.R. Patterson et al. “Does External Walking Environment Affect Gait Patterns?” In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014, pp. 2981–2984. DOI: 10.1109/EMBC.2014.6944249.

[PB10] J. Perry and J.M. Burnfield. *Gait Analysis: Normal and Pathological Gait*. 2nd. Thorofare, NJ, USA: SLACK Inc., 2010.

[PBv01] M. Pijnappels, M.F. Bobbert, and J.H. van Dieën. “Changes in Walking Pattern Caused by the Possibility of a Tripping Reaction”. In: *Gait & Posture* 14.1 (2001), pp. 11–18. DOI: 10.1016/s0966-6362(01)00110-2.

[Ped+11] F. Pedregosa et al. “scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12.85 (2011), pp. 2825–2830.

[Pha+17] M.H. Pham 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”. In: *Frontiers in Neurology* 8 (2017). DOI: 10.3389/fneur.2017.00457.

[Pic+21] P. Picerno et al. “Wearable Inertial Sensors for Human Movement Analysis: A Five-Year Update”. In: *Expert Review of Medical Devices* 18.sup1 (2021), pp. 79–94. DOI: 10.1080/17434440.2021.1988849.

[PK17] W. Pirker and R. Katzenschlager. “Gait Disorders in Adults and the Elderly: A Clinical Guide”. In: *Wiener klinische Wochenschrift* 129.3-4 (2017), pp. 81–95. DOI: 10.1007/s00508-016-1096-4.

[Pol+20] A.M. Polhemus et al. “Walking-Related Digital Mobility Outcomes as Clinical Trial Endpoint Measures: Protocol for a Scoping Review”. In: *BMJ Open* 10.7 (2020). DOI: 10.1136/bmjopen-2020-038704.

[Pol+21] A. Polhemus et al. “Walking on Common Ground: A Cross-Disciplinary Scoping Review on the Clinical Utility of Digital Mobility Outcomes”. In: *NPJ Digital Medicine* 4.1 (2021), p. 149. DOI: 10.1038/s41746-021-00513-5.

[PR91] D. Podsiadlo and S. Richardson. “The Timed ”Up & Go”: A Test of Basic Functional Mobility for Frail Elderly Persons”. In: *Journal of the American Geriatrics Society* 39.2 (1991), pp. 142–148. DOI: 10.1111/j.1532-5415.1991.tb01616.x.

[PSA20] A. Paraschiv-Ionescu, A. Soltani, and K. Aminian. “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 & Biology Society (EMBC)*. 2020, pp. 4596–4599. DOI: 10.1109/EMBC44109.2020.9176281.

[PV08] P. Prandoni and M. Vetterli. *Signal Processing for Communications*. Taylor and Francis Group, LLC, 2008.

[Qua11] Qualisys AB. *Qualisys Track Manager: User Manual*. 2011.

[Ram+15] A. Rampp et al. “Inertial Sensor-Based Stride Parameter Calculation From Gait Sequences in Geriatric Patients”. In: *IEEE Transactions on Biomedical Engineering* 62.4 (2015), pp. 1089–1097. DOI: 10.1109/TBME.2014.2368211.

[Rav+19] A. Ravizza et al. “Comprehensive Review on Current and Future Regulatory Requirements on Wearable Sensors in Preclinical and Clinical Testing”. In: *Frontiers in Bioengineering and Biotechnology* 7 (2019). DOI: 10.3389/fbioe.2019.00313.

[Ray+21] Y.P. Raykov et al. “Probabilistic Modelling of Gait for Robust Passive Monitoring in Daily Life”. In: *IEEE Journal of Biomedical and Health Informatics* 25.6 (2021), pp. 2293–2304. DOI: 10.1109/JBHI.2020.3037857.

[Reh+19] R.Z.U. Rehman et al. “Selecting Clinically Relevant Gait Characteristics for Classification of Early Parkinson’s Disease: A Comprehensive Machine Learning Approach”. In: *Scientific Reports* 9.1 (2019), p. 17269. DOI: 10.1038/s41598-019-53656-7.

[Rém20] P. Rémy. *Temporal Convolutional Networks for Keras*. https://github.com/philipperemy/keras-tcn. 2020.

[Ric18] J. Richards. “The Comprehensive Textbook of Clinical Biomechanics”. In: 2nd. London, UK: Churchill Livingstone, 2018. Chap. Gait Cycle.

[Ris+15] S.M. Rispens et al. “Identification of Fall Risk Predictors in Daily Life Measurements: Gait Characteristics’ Reliability and Association With Self-reported Fall History”. In: *Neurorehabilitation and Neural Repair* 29.1 (2015), pp. 54–61. DOI: 10.1177/1545968314532031.

[RK21] K. Rácz and R.M. Kiss. “Marker Displacement Data Filtering in Gait Analysis: A Technical Note”. In: *Biomedical Signal Processing and Control* 70 (2021), p. 102974. DOI: 10.1016/j.bspc.2021.102974.

[RM00] J.S. Richman and J.R. Moorman. “Physiological Time-Series Analysis Using Approximate Entropy and Sample Entropy”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 278.6 (2000), H2039–H2049. DOI: 10.1152/ajpheart.2000.278.6.H2039.

[Roc+10] S. Rochat et al. “What is the Relationship Between Fear of Falling and Gait in Well-Functioning Older Persons Aged 65 to 70 Years?” In: *Archives of Physical Medicine and Rehabilitation* 91.6 (2010), pp. 879–884. DOI: 10.1016/j.apmr.2010.03.005.

[Roc+20] L. Rochester et al. “A Roadmap to Inform Development, Validation and Approval of Digital Mobility Outcomes: The Mobilise-D Approach”. In: *Digital Biomarkers* 4.1 (2020), pp. 13–27. DOI: 10.1159/000512513.

[Rom+21] R. Romijnders et al. “Validation of IMU-Based Gait Event Detection During Curved Walking and Turning in Older Adults and Parkinson’s Disease Patients”. In: *Journal of NeuroEngineering and Rehabilitation* 18.1 (2021), pp. 1–10. DOI: 10.1186/s12984-021-00828-0.

[Rom+22] R. Romijnders et al. “A Deep Learning Approach for Gait Event Detection from a Single Shank-Worn IMU: Validation in Healthy and Neurological Cohorts”. In: *Sensors* 22.10 (2022). DOI: 10.3390/s22103859.

[Ros+21] R. Rossanigo et al. “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, pp. 1–6. DOI: 10.1109/MeMeA52024.2021.9478604.

[Rot+21a] N. Roth et al. “An Inertial Sensor-Based Gait Analysis Pipeline for the Assessment of Real-World Stair Ambulation Parameters”. In: *Sensors* 21.19 (2021), p. 6559. DOI: 10.3390/s21196559.

[Rot+21b] N. Roth et al. “Do We Walk Differently at Home? A Context-Aware Gait Analysis System in Continuous Real-World Environments”. In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2021, pp. 1932–1935. DOI: 10.1109/EMBC46164.2021.9630378.

[Rot+21c] N. Roth et al. “Hidden Markov Model Based Stride Segmentation on Unsupervised Free-Living Gait Data in Parkinson’s Disease Patients”. In: *Journal of NeuroEngineering and Rehabilitation* 18.1 (2021), p. 93. DOI: 10.1186/s12984-021-00883-7.

[Rue+10] J. Rueterbories et al. “Methods for Gait Event Detection and Analysis in Ambulatory Systems”. In: *Medical Engineering & Physics* 32.6 (2010), pp. 545–552. DOI: 10.1016/j.medengphy.2010.03.007.

[Sab+05] A.M. Sabatini et al. “Assessment of Walking Features from Foot Inertial Sensing”. In: *IEEE Transactions on Biomedical Engineering* 52.3 (2005), pp. 486–494. DOI: 10.1109/TBME.2004.840727.

[Sad+00] H. Sadeghi et al. “Symmetry and Limb Dominance in Able-Bodied Gait: A Review”. In: *Gait & Posture* 12.1 (2000), pp. 34–45. DOI: 10.1016/S0966-6362(00)00070-9.

[Sal+04] A. Salarian et al. “Gait Assessment in Parkinson’s Disease: Toward an Ambulatory System for Long-Term Monitoring”. In: *IEEE Transactions on Biomedical Engineering* 51.8 (2004), pp. 1434–1443. DOI: 10.1109/TBME.2004.827933.

[Sal+13] A. Salarian et al. “A Novel Approach to Reducing Number of Sensing Units for Wearable Gait Analysis Systems”. In: *IEEE Transactions on Biomedical Engineering* 60.1 (2013), pp. 72–77. DOI: 10.1109/TBME.2012.2223465.

[Sal+19] F. Salis et al. “Multi-sensor integration and data fusion for enhancing gait assessment in and out of the laboratory”. In: *Gait & Posture* 74 (2019), p. 34. DOI: 10.1016/j.gaitpost.2019.07.493.

[Sal+21a] F. Salis et al. “A Method for Gait Events Detection Based on Low Spatial Resolution Pressure Insoles Data”. In: *Journal of Biomechanics* 127 (2021), p. 110687. DOI: 10.1016/j.jbiomech.2021.110687.

[Sal+21b] F. Salis et al. “A Wearable Multi-Sensor System for Real World Gait Analysis”. In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2021, pp. 7020–7023. DOI: 10.1109/EMBC46164.2021.9630392.

[SBM16] F.A. Storm, C.J. Buckley, and C. Mazzà. “Gait Event Detection in Laboratory and Real Life Settings: Accuracy of Ankle and Waist Sensor Based Methods”. In: *Gait & Posture* 50 (2016), pp. 42–46. DOI: 10.1016/j.gaitpost.2016.08.012.

[Sch+17] J.C.M. Schlachetzki et al. “Wearable Sensors Objectively Measure Gait Parameters in Parkinson’s Disease”. In: *PLOS ONE* 12.10 (2017), e0183989. DOI: 10.1371/journal.pone.0183989.

[See+10] L. Seematter-Bagnoud et al. “Vulnerability in High-Functioning Persons Aged 65 to 70 Years: The Importance of the Fear Factor”. In: *Aging Clinical and Experimental Research* 22.3 (2010), pp. 212–218. DOI: 10.1007/BF03324799.

[SG14] J.L. Semmlow and B. Griffel. *Biosignal and Medical Image Processing*. Taylor and Francis Group, LLC, 2014.

[SH93] M. Sun and J.O. Hill. “A Method for Measuring Mechanical Work and Work Efficiency During Human Activities”. In: *Journal of Biomechanics* 26.3 (1993), pp. 229–241. DOI: 10.1016/0021-9290(93)90361-H.

[Sha+20] V.V. Shah et al. “Digital Biomarkers of Mobility in Parkinson’s Disease During Daily Living”. In: *Journal of Parkinson’s Diseaes* 10.3 (2020), pp. 1099–1111. DOI: 10.3233/JPD-201914.

[Shl14] J. Shlens. “A Tutorial on Principal Component Analysis”. In: *CoRR* abs/1404.1100 (2014). arXiv: 1404.1100. URL: http://arxiv.org/abs/1404.1100.

[Sli+21] D. Slijepcevic et al. “Explaining Machine Learning Models for Clinical Gait Analysis”. In: *ACM Trans. Comput. Healthcare* 3.2 (2021). DOI: 10.1145/3474121.

[Sni+07] A.H. Snijders et al. “Neurological Gait Disorders in Elderly People: Clinical Approach and Classification”. In: *The Lancet Neurology* 6.1 (2007), pp. 63–74. DOI: 10.1016/S1474-4422(06)70678-0.

[Sol20] A. Soltani. “Gait in Real World: Validated Algorithm for Gait Periods and Speed Estimation Using a Single Wearable Sensor”. PhD thesis. École polytechnique fédérale de Lausanne, 2020.

[Sri+14] N. Srivastava et al. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. In: *Journal of Machine Learning Research* 15.56 (2014), pp. 1929–1958. URL: http://jmlr.org/papers/v15/srivastava14a.html.

[SRS14] T. Seel, J. Raisch, and T. Schauer. “IMU-Based Joint Angle Measurement for Gait Analysis”. In: *Sensors* 14.4 (2014), pp. 6891–6909. DOI: 10.3390/s140406891.

[SRT22] A. Sivakumar, M. Rickman, and D. Thewlis. “Gait Biomechanics After Proximal Femoral Nailing of Intertrochanteric Fractures”. In: *Journal of Orthopaedic Research* (2022). DOI: https://doi.org/10.1002/jor.25427.

[SS20] L.M. Silva and N. Stergiou. “Biomechanics and Gait Analysis”. In: 1st. London, UK: Academic Press, 2020. Chap. The Basics of Gait Analysis, pp. 225–250.

[SSH21] R.M. Schmidt, F. Schneider, and P. Hennig. “Descending Through a Crowded Valley - Benchmarking Deep Learning Optimizers”. In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by Marina Meila and Tong Zhang. Vol. 139. Proceedings of Machine Learning Research. PMLR, 2021, pp. 9367–9376.

[Sta+90] S.J. Stanhope et al. “Kinematic-Based Technique for Event Time Determination During Gait”. In: *Medical and Biological Engineering and Computing* 28.4 (1990), pp. 355–360. DOI: 10.1007/BF02446154.

[Tak17] K. Takakusaki. “Functional Neuroanatomy for Posture and Gait Control”. In: *Journal of Movement Disorders* 10.1 (2017), pp. 1–17. DOI: 10.14802/jmd.16062.

[TCD14] D. Trojaniello, A. Cereatti, and U. Della Croce. “Accuracy, Sensitivity and Robustness of Five Different Methods for the Estimation of Gait Temporal Parameters Using a Single Inertial Sensor Mounted on the Lower Trunk”. In: *Gait & Posture* 40.4 (2014), pp. 487–492. DOI: 10.1016/j.gaitpost.2014.07.007.

[ter19] B. ter Haar Romeny. “Artificial Intelligence in Medical Imaging”. In: Cham: Springer, 2019. Chap. A Deeper Understanding of Deep Learning. DOI: 10.1007/978-3-319-94878-2_3.

[Tho+18] A.J. Thompson et al. “Diagnosis of Multiple Sclerosis: 2017 Revisions of the McDonald Criteria”. In: *The Lancet Neurology* 17.2 (2018), pp. 162–173. DOI: 10.1016/S1474-4422(17)30470-2.

[Tim+18] C. Timmermans et al. “Dual-Task Walking in Challenging Environments in People with Stroke: Cognitive-Motor Interference and Task Prioritization”. In: *Stroke Research and Treatment* 2018 (2018), pp. 1–8. DOI: 10.1155/2018/7928597.

[TNF03] B. Toro, C. Nester, and P. Farren. “A Review of Observational Gait Assessment in Clinical Practice”. In: *Physiotherapy Theory and Practice* 19 (2003), pp. 137–149. DOI: 10.1080/0959398039221901.

[TR20] M. Topley and J.G. Richards. “A Comparison of Currently Available Optoelectronic Motion Capture Systems”. In: *Journal of Biomechanics* 106 (2020), p. 109820. DOI: 10.1016/j.jbiomech.2020.109820.

[Tro+14] D. Trojaniello 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”. In: *Journal of NeuroEngineering and Rehabilitation* 11.1 (2014), p. 152. DOI: 10.1186/1743-0003-11-152.

[Tur+18] A.M. Turcato et al. “Abnormal Gait Pattern Emerges During Curved Trajectories in High-Functioning Parkinsonian Patients Walking in Line at Normal Speed”. In: *PLOS ONE* 13.5 (2018), pp. 1–26. DOI: 10.1371/journal.pone.0197264.

[Ull+20] M. Ullrich et al. “Detection of Gait From Continuous Inertial Sensor Data Using Harmonic Frequencies”. In: *IEEE Journal of Biomedical and Health Informatics* 24.7 (2020), pp. 1869–1878. DOI: 10.1109/JBHI.2020.2975361.

[van+14] K.S. van Schooten et al. “Toward Ambulatory Balance Assessment: Estimating Variability and Stability from Short Bouts of Gait”. In: *Gait & Posture* 39.2 (2014), pp. 695–699. DOI: 10.1016/j.gaitpost.2013.09.020.

[van+15] K.S. van Schooten et al. “Ambulatory Fall-Risk Assessment: Amount and Quality of Daily-Life Gait Predict Falls in Older Adults”. In: *The Journals of Gerontology: Series A* 70.5 (2015), pp. 608–615. DOI: 10.1093/gerona/glu225.

[van+16a] A. van den Oord et al. *WaveNet: A Generative Model for Raw Audio*. 2016. DOI: 10.48550/ARXIV.1609.03499. URL: https://arxiv.org/abs/1609.03499.

[van+16b] J.M.T. van Uem et al. “Twelve-Week Sensor Assessment in Parkinson’s Disease: Impact on Quality of Life”. In: *Movement Disorders* 31.9 (2016), pp. 1337–1338. DOI: https://doi.org/10.1002/mds.26676.

[Van18] T. Van den Heuvel. “Automated Low-Cost Ultrasound: Improving Antenatal Care in Resource-Limited Settings”. PhD thesis. Radboud Universiteit Nijmegen, 2018.

[VD09] G. Van Rossum and F.L. Drake. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009. ISBN: 1441412697.

[Ver+02] J. Verghese et al. “Abnormality of Gait as a Predictor of Non-Alzheimer’s Dementia”. In: *New England Journal of Medicine* 347.22 (2002), pp. 1761–1768. DOI: 10.1056/NEJMoa020441.

[Vie+10] E. Viehweger et al. “Influence of Clinical and Gait Analysis Experience on Reliability of Observational Gait Analysis (Edinburgh Gait Score Reliability)”. In: *Annals of Physical and Rehabilitation Medicine* 53.9 (2010), pp. 535–546. DOI: 10.1016/j.rehab.2010.09.002.

[Vir+20] P. Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[von+95] H.P. von Schroeder et al. “Gait Parameters Following Stroke: A Practical Assessment”. In: *Journal of Rehabilitation Research and Development* 32.1 (1995), pp. 25–31.

[WA19] W. Wang and P.G. Adamczyk. “Analyzing Gait in the Real World Using Wearable Movement Sensors and Frequently Repeated Movement Paths”. In: *Sensors* 19.8 (2019), p. 1925. DOI: 10.3390/s19081925.

[War+20] E. Warmerdam et al. “Long-Term Unsupervised Mobility Assessment in Movement Disorders”. In: *The Lancet Neurology* 19.5 (2020), pp. 462–470. DOI: 10.1016/S1474-4422(19)30397-7.

[War+21] E. Warmerdam et al. “Proposed Mobility Assessments with Simultaneous Full-Body Inertial Measurement Units and Optical Motion Capture in Healthy Adults and Neurological Patients for Future Validation Studies: Study Protocol”. In: *Sensors* 21.17 (2021), p. 5833. DOI: 10.3390/s21175833.

[War21] E. Warmerdam. “The Influence of the Context on Mobility in Neurological Disorders: A Wearable Technology Approach”. PhD thesis. Christian-Albrechts-Universität zu Kiel, 2021.

[WC95] G. Wu and P.R. Cavanagh. “ISB Recommendations for Standardization in the Reporting of Kinematic Data”. In: *Journal of Biomechanics* 28.10 (1995), pp. 1257–1261. DOI: 10.1016/0021-9290(95)00017-C.

[Wil45] F. Wilcoxon. “Individual Comparisons by Ranking Methods”. In: *Biometrics Bulletin* 1.6 (1945), pp. 80–83. DOI: 10.2307/3001968.

[Wol+85] A. Wolf et al. “Determining Lyapunov exponents from a time series”. In: *Physica D: Nonlinear Phenomena* 16.3 (1985), pp. 285–317. DOI: 10.1016/0167-2789(85)90011-9.

[Wor02] World Health Organization. *Towards a Common Language for Functioning, Disability and Health*. https://cdn.who.int/media/docs/default-source/classification/icf/icfbeginnersguide.pdf. Geneva, CH, 2002.

[Wu+02] G. Wu et al. “ISB Recommendations on Definitions of Joint Coordinate System of Various Joints for the Reporting of Human Joint Motion — Part I: Ankle, Hip, and Spine”. In: *Journal of Biomechanics* 35.4 (2002), pp. 543–548. DOI: 10.1016/S0021-9290(01)00222-6.

[YB93] H.J. Yack and R.C. Berger. “Dynamic Stability in the Elderly: Identifying a Possible Measure”. In: *Journal of Gerontology* 48.5 (1993), pp. M225–M230. DOI: 10.1093/geronj/48.5.M225.

[YK16] F. Yu and V. Koltun. “Multi-Scale Context Aggregation by Dilated Convolutions”. In: *4th International Conference on Learning Representations (ICLR)*. 2016. URL: http://arxiv.org/abs/1511.07122v3.

[YKF17] F. Yu, V. Koltun, and T. Funkhouser. “Dilated Residual Networks”. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, 2017, pp. 636–644. ISBN: 978-1-5386-0457-1. DOI: 10.1109/CVPR.2017.75.

[ZH03] W. Zijlstra and A.L. Hof. “Assessment of Spatio-Temporal Gait Parameters from Trunk Accelerations During Human Walking”. In: *Gait & Posture* 18.2 (2003), pp. 1–10. DOI: 10.1016/S0966-6362(02)00190-X.

[Zha+21] A. Zhang et al. “Dive into Deep Learning”. In: *arXiv preprint arXiv:2106.11342* (2021).

[Zij04] W. Zijlstra. “Assessment of Spatio-Temporal Parameters During Unconstrained Walking”. In: *European Journal of Applied Physiology* 92.1 (2004), pp. 39–44. DOI: 10.1007/s00421-004-1041-5.

A

Mathematical Background

Feature Extraction

The **convolutional layer** computes the convolution [Gér19] between the inputs to layer, $\mathbf{A}^{[\ell-1]}$, and some trainable weights, collectively denoted $\mathbf{W}^{[\ell]}$:

$$\mathbf{Z}^{[\ell]} = \mathbf{W}^{[\ell]} * \mathbf{A}^{[\ell-1]}, \quad \mathbf{A}^{[0]} = \mathbf{X} \tag{A.1}$$

where the superscript $.^{[\ell]}$ denotes the ℓth layer, $*$ the convolution operator. Here, appropriate padding is applied such that the number of time steps, N, remains the same for the inputs and outputs. The results of the convolution are:

$$\mathbf{Z}^{[\ell]} = \begin{bmatrix} z_1[1]^{[\ell]} & z_2[1]^{[\ell]} & \cdots & z_{D^{[\ell]}}[1]^{[\ell]} \\ z_1[2]^{[\ell]} & z_2[2]^{[\ell]} & \cdots & z_{D^{[\ell]}}[2]^{[\ell]} \\ \vdots & \vdots & \ddots & \vdots \\ z_1[N]^{[\ell]} & z_2[N]^{[\ell]} & \cdots & z_{D^{[\ell]}}[N]^{[\ell]} \end{bmatrix} \tag{A.2}$$

with $D^{[\ell]}$ the number channels, or the number of kernels, used in the ℓth layer, and:

$$z_{d'}[n]^{[\ell]} = \sum_{d=1}^{D^{[\ell-1]}} \sum_{k=-\lfloor \frac{K}{2} \rfloor}^{\lfloor \frac{K}{2} \rfloor} w_d[k]^{[\ell]} \, a_d[n-k]^{[\ell-1]}, \quad d' = 1, \cdots, D^{[\ell]} \tag{A.3}$$

with K the kernel size, that typically is an odd-valued integer.

The results of the convolution are put through a **batch normalization layer** [IS15]. Given a batch

of M (training) examples, $\left\{\mathbf{Z}^{[\ell](m)}\right\}_{m=1}^{M}$, the batch mean is computed:

$$\bar{\mathbf{Z}}^{[\ell]} = \begin{bmatrix} \bar{z}_1[1]^{[\ell]} & \bar{z}_2[1]^{[\ell]} & \cdots & \bar{z}_{D^{[\ell]}}[1]^{[\ell]} \\ \bar{z}_1[2]^{[\ell]} & \bar{z}_2[2]^{[\ell]} & \cdots & \bar{z}_{D^{[\ell]}}[2]^{[\ell]} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{z}_1[N]^{[\ell]} & \bar{z}_2[N]^{[\ell]} & \cdots & \bar{z}_{D^{[\ell]}}[N]^{[\ell]} \end{bmatrix} \tag{A.4}$$

with $\bar{z}_{d'}[n]^{[\ell]} = \frac{1}{M}\sum_{m=1}^{M} z_{d'}[n]^{[\ell](m)}$, and likewise the batch variance is computed:

$$\mathbf{S}^{[\ell]} = \begin{bmatrix} s_1[1]^{[\ell]} & s_2[1]^{[\ell]} & \cdots & s_{D^{[\ell]}}[1]^{[\ell]} \\ s_1[2]^{[\ell]} & s_2[2]^{[\ell]} & \cdots & s_{D^{[\ell]}}[2]^{[\ell]} \\ \vdots & \vdots & \ddots & \vdots \\ s_1[N]^{[\ell]} & s_2[N]^{[\ell]} & \cdots & s_{D^{[\ell]}}[N]^{[\ell]} \end{bmatrix} \tag{A.5}$$

with $s_{d'}[n]^{[\ell]} = \frac{1}{M}\sum_{m=1}^{M}\left(z_{d'}[n]^{[\ell](m)} - \bar{z}_{d'}[n]^{[\ell]}\right)^2$. These are then used to normalize each of the M examples acoording to:

$$\mathbf{Z}_{\text{norm}}^{[\ell](m)} = \begin{bmatrix} z_{\text{norm},1}[1]^{[\ell](m)} & z_{\text{norm},2}[1]^{[\ell](m)} & \cdots & z_{\text{norm},D^{[\ell]}}[1]^{[\ell](m)} \\ z_{\text{norm},1}[2]^{[\ell](m)} & z_{\text{norm},2}[2]^{[\ell](m)} & \cdots & z_{\text{norm},D^{[\ell](m)}}[2]^{[\ell]} \\ \vdots & \vdots & \ddots & \vdots \\ z_{\text{norm},1}[N]^{[\ell](m)} & z_{\text{norm},2}[N]^{[\ell](m)} & \cdots & z_{\text{norm},D^{[\ell](m)}}[N]^{[\ell]} \end{bmatrix} \tag{A.6}$$

with $z_{\text{norm},d'}[n]^{[\ell](m)} = \left(z_{d'}[n]^{[\ell](m)} - \bar{z}_{d'}[n]^{[\ell]}\right)/\sqrt{(s_{d'}[n]^{[\ell]} + \epsilon)}$ where the ϵ is added for numerical stability just in case the $s_{d'}[n]^{[\ell]}$ turns out to be 0 in some estimates. Now every example has mean 0 and variance 1. However, sometimes it is not desired to have mean 0 and variance, but a different distribution would make more snse, and therefore the following is computed:

$$\tilde{\mathbf{Z}}^{[\ell](m)} = \begin{bmatrix} \tilde{z}_1[1]^{[\ell](m)} & \tilde{z}_2[1]^{[\ell](m)} & \cdots & \tilde{z}_{D^{[\ell]}}[1]^{[\ell](m)} \\ \tilde{z}_1[2]^{[\ell](m)} & \tilde{z}_2[2]^{[\ell](m)} & \cdots & \tilde{z}_{D^{[\ell]}}[2]^{[\ell](m)} \\ \vdots & \vdots & \ddots & \cdots \\ \tilde{z}_1[N]^{[\ell](m)} & \tilde{z}_2[N]^{[\ell](m)} & \cdots & \tilde{z}_{D^{[\ell]}}[N]^{[\ell](m)} \end{bmatrix} \tag{A.7}$$

with $\tilde{z}_{d'}[n]^{[\ell](m)} = \gamma^{[\ell]} z_{\text{norm},d'}[n]^{[\ell](m)} + \beta^{[\ell]}$, where $\gamma^{[\ell]}$ and $\beta^{[\ell]}$ are two more learnable parameters of the model.

The outputs of the batch normalization layer are then fed into a **non-linear activation layer**, that was modelled with a ReLU activation function:

$$\mathbf{A}^{[\ell](m)} = f^{[\ell]}(\tilde{\mathbf{Z}}^{[\ell](m)}) = \begin{bmatrix} a_1[1]^{[\ell](m)} & a_2[1]^{[\ell](m)} & \cdots & a_{D^{[\ell]}}[1]^{[\ell](m)} \\ a_1[2]^{[\ell](m)} & a_2[2]^{[\ell](m)} & \cdots & a_{D^{[\ell]}}[2]^{[\ell](m)} \\ \vdots & \vdots & \ddots & \cdots \\ a_1[N]^{[\ell](m)} & a_2[N]^{[\ell](m)} & \cdots & a_{D^{[\ell]}}[N]^{[\ell](m)} \end{bmatrix} \tag{A.8}$$

with $a_{d'}[n]^{[\ell](m)} = \max(0, \tilde{z}_{d'}[m]^{[\ell](m)})$.

After the non-linear activation function, the data are subject to **dropout** [Sri+14] that randomly "drops out" or deactivates some features, such that the next layer cannot rely on any single feature. Throughout training, on each iteration, standard dropout consists of zeroing out some fraction of the nodes in each layer before calculating the subsequent layer [Zha+21].

Optionally, a **residual connection** is added [He+16; Liu+19]. The residual connection adds the inputs to the residual block to the outputs of the last dropout layer (Figure 2.8a). In case the number of channels between inputs and outputs do not match, then the residual connection consists of a 1×1 convolution, otherwise the inputs are forwarded as they are.

Classification

Now, the outputs from the last residual block, $\mathbf{A}^{[L](m)}$, are considered. The outputs are passed through a **dense** layer that computes a weighted sum of the inputs, for each time step n:

$$\begin{aligned}\mathbf{Z}^{[L+1](m)} &= \mathbf{W}^{[L+1]}\mathbf{A}^{[L](m)} + \mathbf{B}^{[L+1]} \\ &= \begin{bmatrix} z_1[1]^{[L+1](m)} & z_2[1]^{[L+1](m)} & \cdots & z_C[1]^{[L+1](m)} \\ z_1[2]^{[L+1](m)} & z_2[2]^{[L+1](m)} & \cdots & z_C[2]^{[L+1](m)} \\ \vdots & \vdots & \ddots & \vdots \\ z_1[N]^{[L+1](m)} & z_2[N]^{[L+1](m)} & \cdots & z_C[N]^{[L+1](m)} \end{bmatrix}\end{aligned} \tag{A.9}$$

with $z_c[n]^{[L+1](m)} = \sum_{d=1}^{D^{[L]}} w_{d,c}^{[L+1]} a_d[n]^{[L](m)} + b_c^{[L+1]}$ with $c = 1, \cdots, C$ and C is the number of gait events classes.

Multi-Output Model

For the study from Chapter 4 the gait events, in this case the ICs and FCs of a single side, are modelled as separate outputs, thus C = 1 for each of the outputs. The outputs are passed through a sigmoid activation function to give the predictions:

$$\hat{\mathbf{y}}_q^{(m)} = \begin{bmatrix} \hat{y}_q[1]^{(m)} \\ \hat{y}_q[2]^{(m)} \\ \vdots \\ \hat{y}_q[N]^{(m)} \end{bmatrix} = \begin{bmatrix} \sigma\left(z_1[1]^{[L+1](m)}\right) \\ \sigma\left(z_1[2]^{[L+1](m)}\right) \\ \vdots \\ \sigma\left(z_1[N]^{[L+1](m)}\right) \end{bmatrix}, \quad q \in \{\mathrm{IC}, \mathrm{FC}\} \tag{A.10}$$

where $\hat{y}_q[n]^{(m)} = \sigma\left(z_1[n]^{[L+1](m)}\right) = 1/\left(1 + \exp\left(-z_1[n]^{[L+1](m)}\right)\right)$.

The problem is cast as a regression problem, and the difference between the predictions and labels is calculated by means of a weighted mean squared error (MSE). For example, the loss for

the mth example and the qth class is:

$$\ell_q^{(m)} = \frac{1}{N}\sum_{n=1}^{N}\left(y_q[n]^{(m)} + k_q\right)\left(\hat{y}_q[n]^{(m)} - y_q[n]^{(m)}\right)^2, \quad q \in \{\mathsf{IC}, \mathsf{FC}\} \tag{A.11}$$

with k_q a relatively small number (e.g., $k_q \approx 0.01$). The overall loss for the model is then the average loss over all M example and ICs and FCs:

$$\ell = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{2}\left(\ell_{\mathsf{IC}}^{(m)} + \ell_{\mathsf{FC}}^{(m)}\right) \tag{A.12}$$

Single-Output Model

For the study from Chapter 5 the ICs and FCs from both sides are considered at the same time. Hence, taking also into account the null class, $C = 5$. The outputs of the dense layer are passed through a softmax activation layer to give the predictions:

$$\hat{\mathbf{Y}}^{(m)} = \begin{bmatrix} \hat{y}_0[1]^{(m)} & \hat{y}_1[1]^{(m)} & \hat{y}_2[1]^{(m)} & \cdots & \hat{y}_{C-1}[1]^{(m)} \\ \hat{y}_0[1]^{(m)} & \hat{y}_1[2]^{(m)} & \hat{y}_2[2]^{(m)} & \cdots & \hat{y}_{C-1}[2]^{(m)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{y}_0[1]^{(m)} & \hat{y}_1[N]^{(m)} & \hat{y}_2[N]^{(m)} & \cdots & \hat{y}_{C-1}[N]^{(m)} \end{bmatrix} \tag{A.13}$$

with:

$$\hat{y}_{c-1}[n]^{(m)} = \frac{e^{z_c[n]^{[L+1](m)}}}{\sum_{c=1}^{C} e^{z_c[n]^{[L+1](m)}}} \tag{A.14}$$

The problem is cast as a multi-class classification problem, and the difference between the predictions and labels is calculated by means of a weighted categorical crossentropy. For example, the loss for the mth examples is:

$$\ell^{(m)} = \frac{1}{N}\sum_{n=1}^{N}\sum_{c'=0}^{C-1} k_{c'}\; y_{c'}[n]^{(m)} \log\left(\hat{y}_{c'}[n]^{(m)}\right) \tag{A.15}$$

where the index notation is switched from $c = 1, \cdots, C$ to $c' = 0, \cdots, C-1$, and

$$k_{c'} = \begin{cases} 1 & \text{for } c' > 0 \\ 1/f_{\mathsf{s}} & \text{otherwise} \end{cases} \tag{A.16}$$

such that more emphasis is put on correctly detecting gait events ($c' > 0$) than on the null class.

The overall loss for the model is the average loss over all M examples:

$$\ell = \frac{1}{M}\sum_{m=1}^{M}\ell^{(m)} \tag{A.17}$$