POLITECNICO DI TORINO

Master of Science's Degree in Biomedical Engineering

Master of Science's Degree Thesis

Real world mobility assessment with smartphone: validation with Mobilise-D algorithm pipeline and development of a smartphone location recognition framework

Supervisors

Prof. Andrea CEREATTI

Ing. Paolo TASCA

Candidate

Giorgio TRENTADUE

March 2025

Abstract

Human gait is a key health marker, offering critical insights into real-world mobility. Among research-grade dedicated hardware, wearable inertial measurement units (IMUs) are the most widely used for free-living assessments. In the European research initiative Mobilise-D, which aims to develop validated digital mobility outcomes (DMOs) for health monitoring, IMUs serve as the primary sensing technology, enabling the computation of high-accuracy DMOs through a dedicated computational pipeline. Smartphones with built-in inertial sensors offer a cheaper and ubiquitous alternative to IMUs, but differences in positioning and metrological characteristics require re-validation of IMU-based algorithms for phone-derived data. The first objective of this thesis is to assess whether the python implementation of the pipeline, MobGap, originally optimized for lower back-mounted IMU, can achieve comparable results when applied to phone-derived inertial data from the same placement. MobGap assumes that input data is recorded at the lower-back level. Therefore, the second objective focuses on smartphone location recognition with machine learning, including five non-lower back positions. Data were collected at Politecnico di Torino and University of Sheffield. Data collected in Turin included 15 subjects (9 males, 6 females, 23–34 yo) performing in-lab walking tasks. Data collected in Sheffield included 15 subjects (10 males, 5 females, 22–57 yo) performing both structured walking tasks and 2.5h free-living walking sessions. All subjects wore a Samsung Galaxy A34 on the lower back, alongside the INDIP multi-sensor system (pressure insoles + IMUs). In the experiments run at Sheffield, subjects also carried five smartphones in different fixed locations to replicate real-world behaviour. MobGap algorithms were applied to phone derived inertial data and validated against INDIP references. Block-by-block and full pipeline evaluations were conducted. Intermediate DMO such as cadence and stride length were computed using INDIP reference gait sequence and initial contacts. For smartphone location recognition, five machine learning models were compared to distinguish between the six locations. For the first objective, gait sequence detection showed high sensitivity (99.75% in-lab, 98.75% free-living). Initial contact detection achieved 90.4% precision, and 84.3% recall in lab conditions, slightly declining in free-living settings (81.7% precision, 81.9% recall). Cadence estimation had a mean absolute error of 2.12 steps/min in lab conditions, increasing to 5.65 steps/min in free-living conditions. Stride length estimation had a mean absolute error of 14.37 cm in lab conditions, improving to 10.7 cm in free-living conditions. Walking speed, a key metric for Mobilise-D, had an in-lab mean absolute error of 0.13 m/s, decreasing to 0.09 m/s in free-living conditions. For the second objective, the XGBoost model performed best in recognizing smartphone location. While

overall classification across all six positions was moderate (balanced accuracy: 74%), it achieved high accuracy for the lower-back location (96%), demonstrating strong recognition for this placement. These findings suggest that smartphone data can provide comparable gait analysis performance to dedicated IMUs. While lower-back detection was highly accurate, distinguishing other placements (e.g., pocket, hand-held, shoulder bag) remains challenging. Improving this classification could enable MobGap's future application to alternative device locations.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Professor Andrea Cereatti. Without his guidance and support, none of this would have been possible. His expertise and encouragement have been invaluable throughout this journey.

I am also sincerely grateful to Eng. Paolo Tasca, who has closely assisted me throughout this experience. His constant guidance, precision, and unwavering support have played a fundamental role in shaping my work, and I deeply appreciate his presence and dedication.

A heartfelt thank you goes to all the members of the research group in Turin, who have made me feel part of a team where collaboration and mutual support are always present. Their dedication and willingness to share knowledge have made this journey even more enriching.

Returning to Professor Cereatti, I want to express my gratitude once again, not only for his supervision but also for the incredible opportunity he created for me by enabling my research stay at the University of Sheffield.

At Sheffield, I am particularly grateful to Professor Vita Lanfranchi, who made me feel welcome from the very first day. Her kindness and availability have made my time there truly special. Additionally, I would like to thank the entire Department of Computer Science at the University of Sheffield for providing such an innovative and dynamic working environment, which has been an inspiring place to learn and grow.

Finally, I extend my sincere thanks to all the participants in the data collection campaign. Their contribution was essential for gathering the fundamental data that made this study possible.

Table of Contents

List of Tables

List of Figures

Acronyms

DMO
Digital Mobility Outcome

GSD
Gait Sequence Detection

ICD
Initial Contact Detection

CAD
Cadence

SL
Stride Length

WS
Walking Speed

MAE
Mean Absolute Error

ICC
Intra-Class Correlation Coefficient

LoA
Limits of Agreement

IMU
Inertial Measurement Units

MIMU

Magneto Inertial Measurement Units

SLR

Smartphone Location Recognition

LB

Lower Back

BP

Back Pocket

FP

Front Pocket

SB

Shoulder Bag

CP

Coat Pocket

H

Handheld

IC

Initial Contact

Chapter 1

Introduction

1.1 General introduction

The analysis of human gait is a multidisciplinary field with applications in clinical assessment, rehabilitation, sports science, and emerging technologies such as virtual and augmented reality [1]. The integration of data science into gait biomechanics has significantly improved the ability to detect and monitor movement disorders. Machine learning and statistical modeling allow the identification of subtle gait abnormalities that may not be evident through traditional biomechanical methods, facilitating personalized treatment strategies and improving clinical decision-making [2]. Additionally, gait analysis is increasingly recognized for its potential in population health monitoring, enabling large-scale mobility assessments that support early diagnosis and intervention in conditions affecting motor function.

Historically, gait analysis has relied on laboratory-based techniques (such as stereophotogrammetry), which is considered the gold standard for capturing kinematic and spatiotemporal parameters [3]. However, its dependency on controlled laboratory settings, high costs, and complex calibration procedures limit its applicability in real-world scenarios. Giannouli et al. [4] highlighted the distinction between gait capacity, measured in controlled environments, and gait performance, which reflects natural mobility in daily life. This discrepancy underscores the need for wearable sensors that enable continuous and unobtrusive gait monitoring outside the laboratory.

The introduction of Inertial Measurement Units (IMUs) has addressed many of these challenges by enabling gait analysis in free-living conditions. IMUs provide a scalable and accessible solution, allowing continuous monitoring without the constraints of a laboratory. These devices have proven to be valuable tools in rehabilitation, elderly mobility assessment, and sports performance monitoring. Despite their advantages, dedicated IMUs still require intentional placement on the

body, precise calibration, and user compliance, which can limit their widespread adoption for large-scale mobility monitoring.

A natural evolution in gait analysis is the integration of smartphones, which are equipped with built-in inertial sensors and are already widely used in daily life. Unlike dedicated IMUs, smartphones offer a cost-effective and scalable alternative for gait assessment, as they eliminate the need for additional hardware while providing a seamless user experience. This transition enhances the accessibility of gait monitoring, making it feasible for large-scale, long-term mobility tracking in both clinical and non-clinical populations. Smartphones provide a viable platform for gait analysis, as demonstrated by previous studies [5] [6], offering robust data collection beyond controlled environments. Research by Bayat et al. [7] laid the foundation for smartphone-based activity recognition, while more recent work by Olsen et al. [8] validated smartphone-derived spatiotemporal gait parameters against gold-standard methodologies, reinforcing their clinical and real-world applicability.

Among these approaches, the Mobilise-D project represents a major step forward. As a large-scale European research initiative aimed to monitor the daily life gait of people with various mobility problems, Mobilise-D aims to develop and validate digital mobility assessment tools using wearable devices. One of the main outputs of the project is represented by a alorithm pipeline for the assessment of digital mobility outcomes from data recorded by a dedicated lower-back IMU [9], whose implementation has been recently released as the open-source Python package MobGap [10] [9] [11]. By exploiting the pipeline and the data collected throughout the project, Mobilise-D seeks to establish gait analysis as a clinically relevant biomarker, facilitating continuous, real-world monitoring of mobility impairments in various populations, including older adults and patients with chronic conditions. By integrating data-driven methodologies and standardized frameworks, Mobilise-D bridges the gap between controlled laboratory assessments and real-world applications, advancing the role of digital health in personalized medicine.

1.2 Challenges and objectives

The analysis of human gait using dedicated IMUs has been extensively developed for dedicated wearable sensors. These sensors are typically designed for biomechanical assessments, needing precise positioning and secure attachment to the body. However, the growing ubiquity of smartphones, equipped with embedded IMUs, offers a convenient and scalable alternative for gait analysis, thanks to their lower cost and widespread adoption. Despite their potential, smartphone-embedded IMUs exhibit key differences compared to dedicated sensors in terms of metrological properties, placement variability, and attachment methods. These factors can significantly

impact signal quality and gait parameters estimation. Consequently, rigorous validation is necessary to determine whether gait analysis algorithms originally designed for dedicated IMUs can be reliably applied to smartphone-acquired inertial data. The first objective of this thesis is the validation of the Mobilise-D algorithm pipeline, originally validated for data recorded by a waist-worn dedicated IMU, using smartphone-acquired inertial data in both structured environments and real-world settings. This thesis aims to assess the feasibility and accuracy of applying the computational pipeline to smartphone data across different contexts. The validation process involves evaluating key gait parameters estimated by the pipeline, ensuring that results align with those obtained from a wearable gold-standard measurement system. Gait analysis algorithms, including the Mobilise-D pipeline, are typically designed for specific sensor placements on the body. Smartphone location during data collection significantly impacts signal orientation, feature extraction, and ultimately, the accuracy of gait parameter estimation. Since users carry their smartphones in various locations (e.g., trouser pockets, hand, or jacket pockets), identifying the device's placement is essential for ensuring reliable gait analysis and accurate movement interpretation.

To address this, the second objective of this study is to develop a machine learning model for recognizing smartphone location during walking. Given that the Mobilise-D algorithm pipeline is optimized for lower-back IMU data, the classifier will determine whether smartphone-acquired signals can be reliably processed using the same methodology. This classification is particularly important in free-living conditions, ensuring that gait analysis is applied under appropriate circumstances.

In the following, the key challenges of this study are outlined:

- **Experimental Protocol Design**: Developing a structured protocol to investigate gait under various walking conditions, ensuring diverse and representative data collection.
- **Data Preprocessing**: Aligning and standardizing inertial signals from smartphones with those from the reference system to ensure compatibility for the validation process.
- **Integration with the Mobilise-D pipeline**: Adapting and interfacing with the MobGap Python package for correct application of the Mobilise-D pipeline.
- **Multiclass Classification Complexity**: Addressing the inherent difficulties in differentiating smartphone placements solely from inertial data, requiring effective feature extraction and model optimization.

The application of smartphone-based inertial sensors for gait analysis is an expanding research area with significant potential for clinical and digital health

applications. However, ensuring methodological rigor while maintaining practical usability is essential to validate their effectiveness. This study addresses key challenges in leveraging smartphone-acquired inertial data for accurate gait analysis by adapting methodologies originally designed for dedicated IMUs. By systematically assessing the feasibility of using smartphones within the Mobilise-D framework and developing a robust classification model for smartphone location, this research contributes to the advancement of more adaptable and scalable gait analysis techniques.

Beyond validating existing methodologies, this work aims to enhance the generalizability of gait analysis algorithms to accommodate real-world smartphone usage. While traditional IMU-based approaches rely on standardized placements, smartphones introduce variability in positioning that must be accounted for to ensure reliable movement assessment. The development of classification models capable of identifying smartphone placement is a crucial step toward optimizing data interpretation and refining digital mobility assessment tools. As smartphone-based monitoring becomes more prevalent, future advancements will require position-aware algorithms that adjust analytical methods according to device placement, further expanding the applicability of gait analysis in free-living conditions.

1.3 Thesis outline

The structure of this thesis is organized as follows:

- **Chapter 1 - Introduction:**
  This chapter introduces the context of gait analysis, its clinical and technological applications, and the transition from traditional methods to smartphone-based systems. It also outlines the key challenges and objectives of the thesis.

- **Chapter 2 - Background:**
  This chapter provides foundational knowledge about gait, including basic definitions, the gait cycle, and its phases. It reviews traditional and modern gait analysis methods, focusing on spatio-temporal parameter estimation using dedicated IMUs and smartphones. It also discusses the opportunities and challenges of gait analysis in free-living conditions, emphasizing the role of inertial sensors and smartphone positioning.

- **Chapter 3 - Experimental Setup:**
  This chapter describes the components of the reference system used for data acquisition, including IMUs, pressure insoles and distance sensors. It also details the specifications of the Samsung Galaxy A34 smartphone, which serves as the primary data collection device.

- **Chapter 4 - Data Collection:**
  This chapter outlines the experimental protocols for both in-lab and free-living gait data collection. It covers the procedures for data preparation, ensuring consistency and quality across different environments.

- **Chapter 5 - Methods:**
  This chapter details the methodologies employed for validating Mobilise-D algorithm pipeline with smartphone data and developing the machine learning model for smartphone location recognition. This chapter delves into the various steps of the Mobilise-D pipeline and its Python implementation (MobGap), detailing methods for gait sequence detection, initial contact detection, cadence estimation, and stride length estimation. Additionally, the chapter outlines the processes undertaken for developing the smartphone location classifier, including dataset construction, feature extraction, algorithm development, and testing.

- **Chapter 6 - Results:**
  This chapter presents the results of the pipeline validation and the smartphone location recognition model, describing the performance and accuracy of each method.

- **Chapter 7 - Discussion:**
  This chapter analyzes the results in the context of existing literature, discussing the implications, limitations, and potential improvements of the study.

- **Chapter 8 - Conclusions:**
  The final chapter summarizes the key findings of the thesis, highlighting the contributions to the field of gait analysis and suggesting directions for future research.

Chapter 2

Background

2.1 Gait

2.1.1 Basic notions

Gait refers to the pattern of movement of the limbs during locomotion, particularly walking. It is a highly coordinated and repetitive process that enables humans to move efficiently from one place to another while maintaining balance and minimizing energy expenditure. The biomechanics of gait involve the interaction of multiple body segments, joints, and muscles, all working in unison to achieve smooth and stable motion.

According to Perry [12], the human body during gait can be divided into two functionally distinct units:

- **The Locomotive Unit**: This comprises the pelvis and lower limbs, including 11 joints whose movements are controlled by 57 muscles. The bone segments (pelvis, thigh, shank, foot, and toes) alternately support the body and propel it forward, functioning as levers. The locomotive unit serves four main purposes:

  1. **Propulsion**: Driving the body forward.
  2. **Vertical Stability**: Maintaining an upright posture.
  3. **Shock Absorption**: Mitigating impact forces at each step.
  4. **Energy Conservation**: Reducing muscular energy expenditure through efficient biomechanical interactions .

  Additionally, the locomotive unit is responsible for critical motor control functions. The central nervous system integrates feedback from sensory organs to coordinate muscle activity, ensuring correct joint movements and balance adjustments during each phase of the gait cycle. Proper functioning of this

unit helps maintain foot trajectory control, mechanical energy generation for forward propulsion, and energy absorption to moderate body speed [13].

- **The Passenger Unit (Head, Arms, Trunk - HAT)**: Comprising approximately 70% of the body's weight, the passenger unit plays a vital role in sustaining postural stability. While this unit contributes minimally to the active propulsion of movement due to the inherent efficiency of healthy gait mechanics, its proper alignment over the locomotive unit is essential. Misalignment can significantly alter muscle activity in the lower limbs, affecting overall gait stability and efficiency. Moreover, stabilizing the head and upper trunk is crucial for maintaining balance, as it ensures the steadiness of the visual and vestibular systems—key components for spatial orientation and the adaptation of motor strategies during locomotion.

  The passenger unit also aids in corrective postural movements through the swinging of the arms and rotation of the shoulders. This motion counterbalances the rhythmic accelerations and decelerations of the trunk caused by leg movements, reducing rotational forces and enhancing overall gait stability. Variations in arm swing and upper body movement are closely tied to changes in walking speed and balance demands, highlighting the dynamic role of the passenger unit in gait control [13].

The interplay between these units creates a complex, three-dimensional motion framework essential for efficient and stable gait [12].

Understanding these basic notions of gait is fundamental for analyzing walking patterns, identifying abnormalities, and developing interventions for improving locomotor function in both healthy and clinical populations [13]

2.1.2 The gait cycle

The gait cycle or stride cycle refers to the sequence of events that occur from the initial contact of one foot with the ground to the subsequent contact of the same foot. This cyclical process is essential for efficient human locomotion and is typically divided into two primary phases: the stance phase and the swing phase.

- **Stance Phase**: This phase constitutes approximately 60% of the gait cycle and begins with the initial contact (heel strike) and ends when the foot leaves the ground (toe-off). The stance phase provides stability and support as the body moves forward. During this phase, there are two periods of double support, where both feet are in contact with the ground. These periods are critical for maintaining balance, especially during slower walking speeds. As walking speed increases, the duration of double support decreases, eventually

disappearing during running, which introduces a "flight phase" where neither foot is in contact with the ground [14].

- **Swing Phase**: Making up the remaining 40% of the gait cycle, the swing phase begins when the foot lifts off the ground and ends with the next initial contact. This phase is crucial for repositioning the limb in preparation for the next step.

To further analyze gait, three primary approaches are used:

1. **Foot-to-Floor Contact Events**: This method segments the gait cycle based on the alternating contact of the feet with the ground. It highlights the periods of single and double support, where either one or both feet are in contact with the ground.

2. **Temporal and Spatial Features**: This approach focuses on time-related (e.g., stride duration) and distance-related (e.g., step length) aspects of the gait cycle, providing a quantitative perspective.

3. **Functional Phases**: The gait cycle is divided into functional phases based on key events such as initial contact, toe-off, and mid-stance. These phases describe the biomechanical roles of the limbs during walking, including weight acceptance, single-limb support, and limb advancement [12].

Understanding the taxonomy of gait allows for precise analysis of walking patterns, aiding in the identification of normal and pathological gait characteristics.

2.1.3 Phases and sub-phases of gait

According to Perry [12], the gait cycle can be hierarchically divided into periods, tasks, and sub-phases, each representing specific motor functions necessary for efficient locomotion. The two main periods are the stance phase and the swing phase, which are further subdivided into functional tasks and sub-phases.

Stance Phase Sub-phases:

1. **Initial Contact** (0-2% of Gait Cycle): The moment the foot first touches the ground, typically with the heel.

2. **Loading Response** (2-10%): The period immediately following initial contact, where the body weight is transferred onto the leading limb, absorbing impact forces.

3. **Mid Stance** (10-30%): The body progresses over the stationary foot, achieving a stable, single-limb support.

4. **Terminal Stance** (30-50%): The heel lifts off as the body continues forward, preparing for the next phase.

5. **Pre-Swing** (50-60%): Also known as the "toe-off" phase, the foot leaves the ground, marking the transition to the swing phase.

Swing Phase Sub-phases:

1. **Initial Swing** (60-73%): The limb accelerates forward, clearing the ground.

2. **Mid Swing** (73-87%): The limb continues its forward motion, aligning with the stance leg.

3. **Terminal Swing** (87-100%): The limb decelerates, preparing for the next initial contact.

These sub-phases collectively fulfill three critical functional tasks:

- **Weight Acceptance**: Involves initial contact and loading response, focusing on absorbing impact and stabilizing the body. This is the most demanding task in the gait cycle, as it requires an abrupt transfer of body weight onto a limb that has just completed its swing phase [14].

- **Single Limb Support**: Comprises mid stance and terminal stance, maintaining balance as the body progresses over one limb.

- **Limb Advancement**: Encompasses pre-swing and all swing phases, repositioning the limb for the next step.

This hierarchical structure provides a comprehensive framework for analyzing gait mechanics, essential for both clinical assessments and biomechanical research [12].

2.2 Gait analysis

Gait analysis is the systematic study of human walking, integrating biomechanical, physiological, and clinical assessments to understand locomotion patterns. It is a pivotal tool in both research and clinical environments, aiding in the diagnosis of gait abnormalities, optimizing rehabilitation strategies, and enhancing athletic performance.

2.2.1 General overview

The formalization of gait analysis as a scientific discipline began in the late 19th and early 20th centuries. The introduction of photographic techniques by Eadweard Muybridge and Étienne-Jules Marey allowed for the visualization of movement sequences, revolutionizing the understanding of gait cycles. Their pioneering work laid the groundwork for motion analysis by capturing detailed images of human locomotion. Figure 2.1 illustrates one of Muybridge's famous photographic sequences, depicting the different phases of a walking cycle.

Figure 2.1: Eadweard Muybridge's photographic sequence depicting human locomotion, illustrating the different phases of the walking cycle.

Following World War II, electromyography became the primary method for studying muscle activity during walking. Jacquelin Perry [12] [15] and David Sutherland [16], prominent figures in the field, significantly advanced gait analysis through the development of instrumented methods and three-dimensional motion analysis. Their contributions established the foundation for modern gait laboratories, enabling precise assessments of locomotor patterns.

The advent of computer technology in the 1970s further propelled gait analysis into the modern era. Computational advancements allowed for more efficient data processing and the integration of complex biomechanical models, making gait analysis more accessible and clinically relevant [17]. However, as highlighted by Sheldon [18], despite these technological advancements, the routine clinical use of gait analysis has faced limitations due to the complexity of data interpretation, high costs, and lengthy testing procedures, which have slowed its widespread adoption in clinical settings.

In the 1990s and early 2000s, the emergence of wearable sensor technologies marked a significant shift in gait analysis. Tao et al. [19] emphasized the development of accelerometers, gyroscopes, and magnetoresistive sensors, which allowed for portable and cost-effective gait assessments outside of traditional laboratory environments. These innovations expanded the applicability of gait analysis to real-world settings, facilitating its use in rehabilitation, sports performance, and continuous health monitoring.

More recently, the integration of data science and machine learning has revolutionized gait biomechanics. Ferber et al. [1] discussed how modern data science methods, have enhanced the ability to analyze large, complex datasets, providing deeper insights into gait patterns and improving clinical decision-making. These approaches enable the identification of subtle gait abnormalities and support personalized treatment planning by leveraging big data collected from multiple sources.

Today, gait analysis encompasses multiple scientific domains:

- **Kinematics**: Examining limb and joint movements using motion capture systems, providing precise data on body segment trajectories [20].
- **Clinical Applications**: Implementing gait analysis in the diagnosis and monitoring of neurological disorders (e.g., Parkinson's disease), musculoskeletal conditions (e.g., post-arthroplasty recovery), and in optimizing rehabilitation protocols [20][21][22].
- **Electromyography**: Assessing muscle activation patterns to evaluate neuromuscular control and coordination [20].
- **Energy Expenditure**: Estimating the metabolic cost of walking, often through indirect calorimetry or wearable sensors [20].

Modern gait analysis utilizes both sophisticated laboratory-based systems and portable, wearable technologies that enable data collection in real-world environments. Clinically, gait analysis is vital for diagnosing conditions like Parkinson's disease, stroke-related impairments, and musculoskeletal disorders. For instance, Morris and Summers [21] demonstrated how gait analysis can clarify abnormalities in stride length-cadence relationships in Parkinsonian gait, offering valuable insights into the effects of levodopa medication on locomotor patterns. Similarly, Temporiti et al. [22] utilized gait analysis to compare functional recovery in patients undergoing bilateral versus unilateral total hip arthroplasty, highlighting its role in tailoring rehabilitation strategies. Whittle [20] emphasized the broader clinical applications of gait analysis, particularly in cerebral palsy and stroke rehabilitation, illustrating its utility in both diagnostic and therapeutic contexts. These examples underscore how gait analysis not only aids in identifying specific gait abnormalities but also

in evaluating the efficacy of interventions, from pharmacological treatments to surgical and rehabilitative procedures.

In summary, mobility assessment is an indispensable tool that bridges multiple disciplines, offering comprehensive insights into human locomotion, supporting clinical decision-making, and advancing the development of assistive technologies.

2.2.2 Spatio-temporal parameters

Spatio-temporal parameters are fundamental metrics in gait analysis, providing quantitative descriptions of the temporal (time-related) and spatial (distance-related) characteristics of walking. These parameters are essential for identifying deviations from normal gait patterns and for evaluating the efficacy of therapeutic interventions in clinical settings [20].

Temporal Parameters:

- **Gait Cycle Duration**: *(Also known as Stride Duration)* The total time taken for a single gait cycle, from the initial contact of one foot to the subsequent contact of the same foot.
- **Stance Time**: The duration a foot remains in contact with the ground during the gait cycle, typically representing about 60% of the cycle.
- **Swing Time**: The period during which the foot is off the ground, making up the remaining 40% of the gait cycle.
- **Double Support Time**: The time when both feet are in contact with the ground, occurring twice in each gait cycle. This parameter decreases with increasing walking speed and disappears during running [23].
- **Cadence**: The number of steps taken per minute, reflecting the rhythm of walking.

Spatial Parameters:

- **Step Length**: The distance between the initial contact point of one foot and the initial contact point of the opposite foot.
- **Stride Length**: The distance covered in one complete gait cycle, typically equal to two step lengths.
- **Step Width**: The lateral distance between the paths of the two feet, indicating stability and balance.
- **Walking Speed**(Gait Velocity): The overall rate of walking, calculated as the product of cadence and step length. It is a critical indicator of functional mobility [23].

Several methodologies have been developed to accurately capture spatio-temporal parameters, each tailored to specific research and clinical contexts. Traditional gait laboratories often employ motion capture systems with reflective markers [24], enabling the tracking of body movements in three dimensions. While these systems provide highly accurate spatial and temporal data, their cost and the necessity for controlled environments limit their widespread use [20].

Another standard method involves the use of force platforms [24], which measure ground reaction forces. These platforms allow precise determination of temporal parameters like stance and swing times and provide insights into dynamic aspects of gait such as propulsion and balance [20].

With the advent of portable technology, wearable IMUs have become increasingly popular. These devices, including accelerometers, gyroscopes, and magnetometers—referred to as Magneto-Inertial Measurement Units (MIMUs) when magnetometers are included—are embedded in wearable sensors and offer a cost-effective, flexible solution for capturing gait data outside of laboratory environments. Tao et al. [19] highlighted the effectiveness of IMUs in real-world settings, marking a significant shift in how gait analysis is conducted.

Similarly, image-based systems like the Microsoft Kinect have introduced new possibilities for gait analysis. Souza and Stemmer [25] demonstrated the use of Kinect sensors to extract kinetic and kinematic parameters through three-dimensional imaging. This technology provides a non-invasive, low-cost alternative for both clinical assessments and biometric applications. Pedobarographic systems, which use pressure-sensitive mats and insoles, offer another valuable approach. These systems measure foot pressure distribution during walking, yielding detailed insights into spatial parameters such as step width and gait symmetry, which are critical for assessing balance and stability [20].

In addition to these technological advancements, normative data studies remain fundamental for contextualizing individual gait assessments within population benchmarks. Oberg and Oberg [23] conducted an extensive analysis of spatio-temporal gait parameters, compiling data from a broad cohort spanning ages 10 to 79. Their research highlighted distinct trends: walking speed and step length were shown to decrease progressively with age, reflecting natural declines in muscle strength, joint flexibility, and overall mobility. Interestingly, cadence remained relatively stable across age groups, suggesting compensatory mechanisms to maintain gait rhythm despite shorter steps. The study also revealed gender differences, with men typically exhibiting longer step and stride lengths, while women demonstrated higher cadence rates. These normative datasets are critical for clinical assessments, enabling practitioners to differentiate between normal age-related gait changes and pathological deviations, thus refining diagnostic accuracy and the personalization of rehabilitation strategies. Understanding and accurately measuring spatio-temporal parameters are crucial for developing effective

rehabilitation protocols and advancing research in human locomotion.

2.2.3 Estimating spatio-temporal parameters from dedicated IMUs

As introduced in Subsection 2.2.2, IMUs have revolutionized the estimation of spatio-temporal parameters in gait analysis, offering a portable and versatile alternative to traditional motion capture systems.

The methodology for utilizing IMUs in gait analysis has evolved significantly, with foundational work by Zijlstra and Hof [26] laying the groundwork. Their approach used accelerometry data from the trunk combined with inverted pendulum models to estimate stride length, a technique that has proven robust across various populations and walking conditions. Subsequent refinements by Zijlstra and Zijlstra [27] further validated this method, emphasizing its reliability.

Della Croce and Mancini [28] provided a comprehensive overview of the methodologies involved in extracting spatio-temporal parameters from IMU data. Sensor placement plays a critical role in data accuracy, with common sites including the lower back, thighs, shanks, and feet—each offering specific advantages depending on the parameter being measured. For example, positioning an IMU on the lower back aligns it closely with the body's center of mass, facilitating reliable data on overall gait dynamics, while sensors on the feet capture detailed foot-ground interactions essential for precise step analysis.

Once data is collected, it undergoes sophisticated signal processing to extract meaningful gait parameters like step length, stride duration, cadence, and walking speed. Techniques such as Kalman filtering, complementary filtering, and sensor fusion algorithms help mitigate noise and correct for sensor drift. Event detection algorithms, including threshold-based detection and zero-crossing methods, identify key gait events like heel strikes and toe-offs, pivotal for accurate parameter estimation.

Several methods have been reported in the literature for estimating spatial parameters from inertial signals. Direct integration involves double integrating the linear acceleration signals to derive displacement. While conceptually straightforward, this method is highly sensitive to noise and drift, often leading to cumulative errors. To address this, corrective strategies such as zero-velocity updates and direct-reverse integration are commonly employed.

Biomechanical models and abstraction techniques, such as the inverted pendulum model, mathematically represent gait dynamics. These models leverage anthropometric data and motion dynamics to estimate parameters like stride length. The work by Zijlstra and Hof [26] remains a cornerstone in this area, and their methodologies continue to be widely used in modern gait analysis.

The application of IMUs in gait analysis has been extensively validated in both

clinical and research settings. Bovonsunthonchai et al. [29] demonstrated the effectiveness of IMUs in capturing spatio-temporal parameters in patients with Parkinson's disease, facilitating the monitoring of disease progression and evaluation of treatment efficacy. Trojaniello et al. [30] further validated IMUs for accurate foot contact detection and parameter estimation in individuals with gait abnormalities. IMUs have also been used to assess gait impairments in stroke survivors, individuals with musculoskeletal disorders, and elderly populations at risk of falls, enabling personalized rehabilitation programs and improving clinical outcomes [31] [32].

Several mobility assessment solutions based on dedicated IMUs are available for research or clinical purposes on the market, such as the DynaPort sensor by McRoberts, the G-Walk sensor by BTS, and others. Despite their many benefits, dedicated IMUs also present several challenges. One of the main drawbacks is the need for specialized software and proprietary tools, which can limit accessibility and flexibility in data processing. Additionally, dedicated IMUs tend to have a relatively high cost, often reaching several hundred euros, making them less feasible for large-scale studies.

Another significant limitation is their operational complexity, which can reduce their usability in real-world applications. These devices often require complex calibration procedures, manual control over data acquisition, and external tools for data transfer and processing. Such requirements increase the time and effort needed for data collection, limiting their practicality outside controlled research or clinical environments.

In light of these challenges, technologies that retain the benefits of IMUs while offering greater accessibility and ease of use have been explored. One promising direction is represented by the use of smartphones, which are equipped with built-in inertial sensors similar to those in dedicated IMUs. The next section will explore how smartphones can be effectively utilized for gait analysis, highlighting their potential to democratize access to this important diagnostic tool.

2.2.4 Estimating spatio-temporal parameters from smartphones

The integration of advanced inertial sensors in smartphones has enabled the extraction of spatio-temporal gait parameters, offering a cost-effective and accessible alternative to dedicated IMUs. Modern smartphones are equipped with accelerometers, gyroscopes, and magnetometers, mirroring the core functionalities of IMUs and allowing for the estimation of spatio-temporal parameters in both clinical and everyday environments.

Previous research has demonstrated that smartphones can achieve accuracy comparable to dedicated IMUs in controlled conditions, as they leverage similar methodologies for extracting spatio-temporal parameters. Several studies have

validated the effectiveness of smartphone-based mobility assessment, employing techniques such as signal integration, frequency-domain analysis, and biomechanical modeling to estimate stride length, walking speed, cadence, and step detection.

Bayat et al. [7] applied direct integration methods, refining displacement estimation through low-pass filtering, and demonstrated that stride length and walking speed can be reliably measured using smartphone accelerometer data. Similarly, Kang et al. [33] utilized gyroscope data and frequency-domain analysis via Fast Fourier Transform (FFT) to enhance step detection, cadence estimation, and walking pattern recognition, even in unconstrained smartphone placements.

Beyond direct signal integration, biomechanical models have been effectively adapted for smartphones to estimate spatio-temporal gait parameters. Serra-Añó et al.[34] employed inverted pendulum models to assess gait cycle duration, step length variability, and gait asymmetry in individuals with Alzheimer's disease, capturing subtle mobility impairments under single-task and dual-task conditions. Olsen et al.[8] validated smartphone-based gait analysis by integrating double inverted pendulum models, double integration, and peak detection techniques, demonstrating that gait parameters such as step length, cadence, and gait speed exhibited moderate to excellent validity when compared to the GAITRite® system. [35]

Machine learning techniques have further enhanced the accuracy of smartphone-derived spatio-temporal parameters. Hannink et al. (2017) [36] utilized deep convolutional neural networks to estimate stride length, step duration, and variability, showcasing the potential of AI-driven methodologies in refining gait assessments. Olsen et al.[8] further demonstrated how machine learning models improve step segmentation and variability detection, making smartphone-based gait analysis even more precise.

Many studies validating smartphone-based methods are conducted in controlled environments, ensuring repeatability and reducing external variability. However, these controlled conditions can differ significantly from real-life scenarios, where factors such as uneven terrain, environmental distractions, and variations in walking behavior may affect gait analysis.

The ability of smartphones to extract spatio-temporal gait parameters in real-world settings offers significant advantages for continuous mobility monitoring. This capability is particularly relevant for assessing patients with Parkinson's disease, stroke, musculoskeletal disorders, and neurodegenerative conditions such as Alzheimer's, where changes in gait parameters can provide early indicators of mobility impairments. By offering an unobtrusive, scalable, and cost-effective solution, smartphones democratize access to spatio-temporal gait analysis, enabling large-scale mobility assessments both in clinical practice and at home.

2.2.5 Gait analysis in free-living conditions: opportunities and challenges

Gait analysis in free-living conditions represents a significant shift from traditional laboratory-based assessments, offering the potential to capture a more authentic representation of an individual's mobility. While controlled environments provide precise measurements, they may not fully reflect the complexities and variabilities of daily life. Free-living gait analysis addresses this gap, allowing continuous monitoring in everyday-life settings and offering valuable insights into real-world functional mobility.

One of the primary opportunities presented by free-living gait analysis is its ability to reveal fluctuations in gait patterns that may be missed in clinical settings. Galperin [37] demonstrated that gait metrics collected through wearable sensors in daily life explained a significant portion of the variability in motor symptom severity among Parkinson's disease patients. These real-world measurements provided complementary information to traditional clinical assessments, highlighting the importance of monitoring mobility in natural environments to better understand disease progression and functional decline.

Similarly, Giannouli [4] and Zijlstra [27] emphasized the distinction between mobility capacity, as measured in laboratory tests, and mobility performance in daily life. Their study showed that laboratory-based gait metrics explained only a small fraction of the variance in real-life mobility behaviors. This suggests that real-world gait performance is influenced by additional factors such as cognitive load, environmental context, and social interactions, underscoring the value of free-living gait analysis for a holistic understanding of functional mobility.

The integration of wearable sensors and smartphones has facilitated the widespread adoption of free-living gait analysis. These technologies enable unobtrusive, continuous monitoring of gait parameters, providing clinicians and researchers with rich datasets that capture the nuances of daily mobility. This data is particularly valuable for tailoring personalized interventions, tracking treatment efficacy, and detecting early signs of mobility impairments.

Despite its many advantages, free-living gait analysis also presents several challenges. The variability introduced by uncontrolled environmental factors, such as uneven terrain, varying lighting conditions, and the presence of obstacles, can affect the accuracy and consistency of gait measurements. Moreover, ensuring data reliability in the face of device placement variability and user compliance remains a critical concern.

Privacy and data security are also paramount considerations in free-living gait analysis. The continuous collection of sensitive personal data necessitates robust data protection measures and adherence to ethical standards to ensure user trust and compliance with regulatory requirements.

Looking forward, the future of gait analysis in free-living conditions lies in the continued advancement of wearable technologies and data analytics. The development of smart textiles, unobtrusive sensors, and real-time feedback systems will further enhance the feasibility and accuracy of free-living gait assessments. Additionally, the integration of gait data with other health metrics, such as heart rate and activity levels, will provide a more comprehensive view of an individual's overall health and well-being. In conclusion, gait analysis in free-living conditions offers unparalleled opportunities to understand human mobility in real-world contexts.

2.3 Use of Inertial Sensors for Free-Living Gait Analysis

2.3.1 Principles of Operation of IMUs

Inertial sensors enable the tracking of an object's motion relative to an inertial reference frame [38]. These sensors directly capture inertial signals, such as accelerations and angular velocities. Through successive integration of these measurements, it is possible to derive linear velocities, linear displacements, and angular displacements. Due to their affordability and sufficient reliability, inertial sensors are widely employed across various fields, ranging from military applications to consumer services. However, this technology is subject to certain limitations, including drift errors and the necessity for frequent calibration. As introduced in Subsection 2.2.2, inertial sensors are often integrated with magnetometers within MIMU.

Accelerometer

Accelerometers measure the proper linear acceleration a_p, which is the vectorial difference between the coordinate acceleration a_c—representing the rate of change of the sensor's velocity—and the acceleration due to gravity, g:

$$a_p = a_c - g \tag{2.1}$$

Thus, the output of an accelerometer does not directly represent the actual acceleration experienced by the sensor but rather the difference between coordinate acceleration and gravitational acceleration. As a result, a one-dimensional accelerometer in free fall, with its positive axis aligned with gravity, will output 0 $\mathrm{m/s^2}$, whereas it will measure $|g| = 9.81$ $\mathrm{m/s^2}$ when stationary.

An accelerometer can be modeled as a second-order spring-mass-damper system, consisting of a proof mass m, an elasticity constant k, and a damping factor b ((Figure 2.2)).

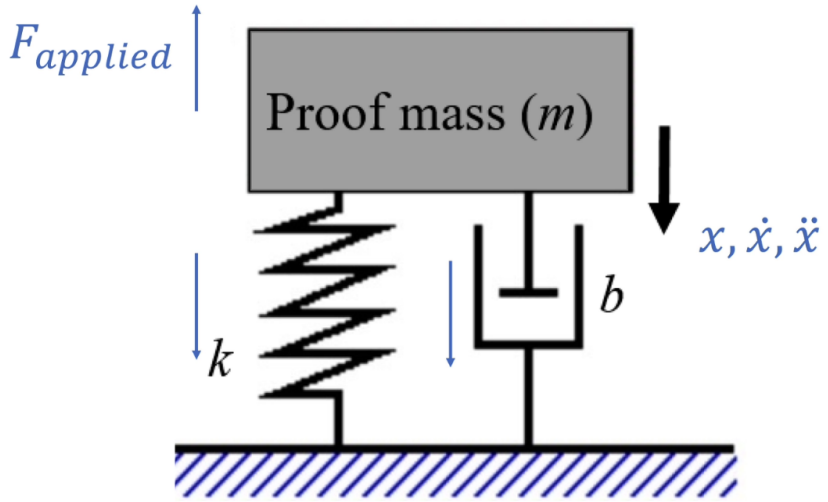


Figure 2.2: Dynamic model of a 1D-accelerometer [39]

When an external force F_{applied} is exerted on the system, the proof mass reacts under the influence of elastic and damping forces, in accordance with Newton's second law:

$$F_{\text{applied}} + F_{\text{spring}} + F_{\text{damper}} = F_{\text{mass}} \tag{2.2}$$

Substituting the expressions for the elastic and damping forces leads to the equation:

$$kx + b\dot{x} + m\ddot{x} = ma_c \tag{2.3}$$

where a_c is the coordinate acceleration. In the absence of gravity, a_c corresponds to the proper acceleration. However, if the accelerometer's sensitivity axis is aligned with gravity, the measured acceleration a_p includes the gravitational component:

$$kx + b\dot{x} + m\ddot{x} = m(a_c - g) = ma_p \tag{2.4}$$

The accelerometer's response to an input acceleration is characterized by its transfer function $H(s)$, which depends on the system's resonance frequency ε_0 and quality factor Q:

$$H(s) = \frac{1}{\varepsilon_0^2 + Qs + s^2} \tag{2.5}$$

where:

$$Q = \frac{m\varepsilon_0}{b}, \quad \varepsilon_0 = \sqrt{\frac{k}{m}} \tag{2.6}$$

The sensitivity of the accelerometer decreases with the square of frequency, from an initial static sensitivity $H(0)$ when subjected to constant acceleration to zero at very high frequencies. As a result, accelerometers are typically designed to operate well below their resonance frequency to ensure reliable measurements. However, increasing the resonance frequency expands the sensing bandwidth at the cost of reducing sensitivity, requiring a trade-off between these parameters in accelerometer design [39].

Transduction mechanism

Several sensing mechanisms can be employed to convert the displacement of the proof mass, induced by input acceleration, into a measurable signal. Among the various transduction principles available, the most commonly used types of accelerometers can be classified as follows:

- **Capacitive:** In capacitive accelerometers, the displacement of the proof mass is first converted into a proportional capacitance change, which is then transformed into an amplified voltage signal. These accelerometers offer high sensitivity, excellent DC performance, low noise, minimal drift, and reduced temperature dependence. However, they are highly susceptible to electromagnetic interference [39].
- **Piezoresistive:** These accelerometers leverage the property of piezoelectric crystals to generate a change in resistivity in response to mechanical stress along a sensitive axis. They also have a simple design, straightforward fabrication process, and easy-to-read output circuits. However, they typically exhibit low sensitivity, bulky structures, and significant sensitivity to temperature variations [39].
- **Piezoelectric:** These accelerometers leverage the property of piezoelectric crystals to generate an electric charge in response to mechanical stress along a sensitive axis. Their main advantages include being self-powered and providing a direct digital output with relatively simple interface circuitry. However, their performance degrades under DC conditions due to charge leakage, and they tend to have larger physical dimensions [39].

Gyroscope

Gyroscopes measure angular velocity along one, two, or three sensitivity axes. Devices capable of measuring angular velocity in all three directions are commonly

referred to as tri-axial gyroscopes, and their measurements are typically expressed in degrees per second (dps) or radians per second (rad/s). When combined with accelerometers, gyroscopes enable various applications that require sensor fusion of integrated inertial sensing [40].

The most widely used gyroscopes operate based on the *Coriolis effect*, also known as vibrating gyroscopes. These devices can be modeled as a *spring-mass-damper system* with two degrees of freedom [39] [40], as illustrated in Figure 2.3.

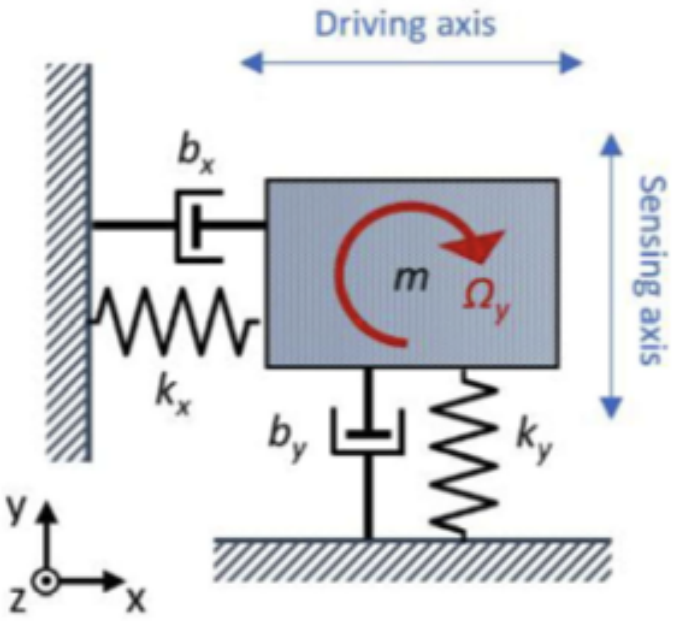


Figure 2.3: Spring-mass-damper system used to model the gyroscope.

According to the Coriolis principle, when an object with mass m undergoes both rotation at an angular velocity $\boldsymbol{\Omega_z}$ and linear motion at velocity $\boldsymbol{v_t}$, a force—referred to as the *Coriolis force*—acts on the object:

$$\boldsymbol{F_c} = 2m\boldsymbol{v_t} \times \boldsymbol{\Omega_z} \tag{2.7}$$

In vibrating fork gyroscopes, a pair of proof masses oscillate with equal amplitude but in opposite directions. When the device is stationary, these tines resonate in anti-phase within the plane of the fork (*drive mode*). Upon rotation, an additional oscillation component appears along the direction orthogonal to the plane, generating a torque that excites a *torsional mode* around the gyroscope stem. The forces acting on the proof mass in both the drive and sensing directions can be defined using Newton's second law [40]:

$$F_x = m\ddot{x} + b_x\dot{x} + k_x x \tag{2.8}$$

$$F_y - 2m\Omega_z\dot{x} = m\ddot{y} + b_y\dot{y} + k_y y \tag{2.9}$$

These equations illustrate that angular velocity can be estimated by analyzing the displacement of the proof mass.

Among the various gyroscope designs, the most widely used solution is based on capacitive sensing, which can be arranged in different configurations depending on complexity levels [41] [42]. The simplest gyroscope structure is shown in Figure 2.4 and consists of:

- A proof mass attached to a substrate.
- Driving electrodes, which induce a constant oscillatory motion along the drive axis.
- Sensing electrodes, which detect the Coriolis force.
- A suspension system composed of four suspension pillars [42] [39].

The combined motion of the driving and sensing electrodes induces a differential capacitance change, which is used to measure the Coriolis force and, consequently, determine the angular velocity.

Unlike accelerometers, which can function as passive components, vibrating fork gyroscopes necessarily require an active element, such as driving electrodes, to sustain the oscillatory motion needed for Coriolis-based angular velocity sensing.

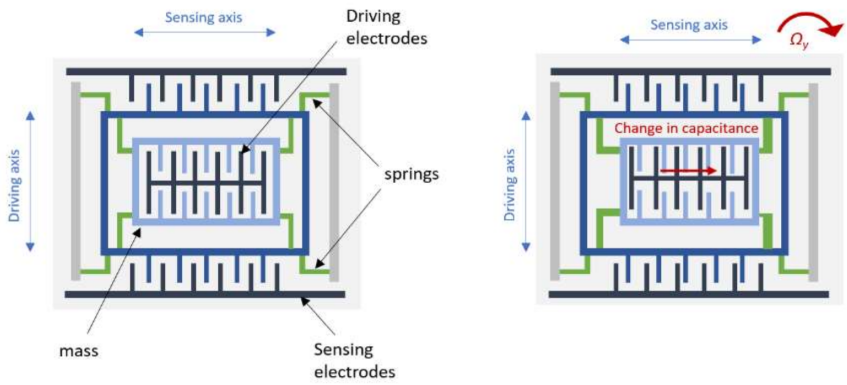


Figure 2.4: Schematic representation of a capacitive gyroscope, illustrating the working principle based on Coriolis force detection through differential capacitance changes.

2.3.2 State of the Art in the Use of Dedicated IMUs for gait assessment

The use of dedicated IMUs in gait analysis has evolved significantly over the past decades, establishing itself as a cornerstone in both clinical and research applications.

Setup and Methods Used

The typical setup for IMU-based gait analysis involves strategically placing sensors on key body segments, such as the lower back, thighs, shanks, and feet. This configuration allows for the capture of detailed kinematic data, including joint angles, stride length, and walking speed. The accuracy of these measurements is enhanced through advanced signal processing techniques, such as Kalman filtering and sensor fusion algorithms [43], which mitigate sensor drift and noise.

Several methodologies have been developed to extract meaningful gait parameters from IMU data. Direct integration methods, while straightforward, often require additional correction techniques like zero-velocity updates to counteract cumulative errors. Biomechanical models, such as the inverted pendulum model, provide a theoretical framework for interpreting gait dynamics and have been widely adopted for their robustness and reliability. More recently, machine learning algorithms have been employed to enhance the accuracy of gait parameter estimation, leveraging large datasets to identify complex patterns in motion data.

During the last six years, significant step forwards have been done in the framework of the Mobilise-D project. The Mobilise-D project was an innovative European initiative aimed at transforming the assessment of human mobility through advanced digital technologies. Funded under the Horizon 2020 program, Mobilise-D developed and validated sensor-based tools for accurately and continuously measuring mobility in individuals with chronic conditions such as Parkinson's disease, multiple sclerosis, and cardiovascular diseases. The ultimate goal was to enhance diagnosis, treatment, and clinical management by providing reliable and standardized data that can support personalized therapeutic decisions and improve patient outcomes.

Within the scope of the technical validation study performed by the Mobilise-D consortium, several studies addressed the challenge of digital moobility assessment in free-living conditions with an IMU mounted on the lower-back. Kluge et al.[44] proposed a consensus-based framework for the digital monitoring of mobility, aiming to standardize the use of IMUs for gait analysis in both clinical and daily life contexts. Their framework addresses challenges related to the reliability and validity of IMU data, ensuring consistent methodologies across different studies.

Mazzà et al.[45] conducted a multicenter validation study for real-world gait monitoring using IMUs. They performed extensive metrological verification of

IMU devices and validated algorithms for calculating digital mobility outcomes in real-life conditions across various clinical populations, including Parkinson's disease and COPD. Their results demonstrated high reliability and validity of IMUs in capturing spatio-temporal gait parameters, highlighting their applicability in diverse clinical scenarios.

Bertuletti et al. [46] introduced an innovative wearable platform combining MIMUs with infrared proximity sensors to measure inter-foot distance. This approach overcomes the limitations of traditional IMUs by improving the accuracy of step stability assessments. Their study reported a mean error of less than 1.5 cm in the estimation of step width, demonstrating the potential of this hybrid system for precise gait analysis.

Salis et al. have progressively developed and validated a multi-sensor system to enhance gait assessment in real-world conditions. In 2014 [47], they introduced a method using low-resolution pressure insoles for gait event detection, providing a cost-effective and portable alternative for integration with IMU-based systems. Building upon this work, in 2019 [48], they presented the INDIP system, a wearable multi-sensor setup combining IMUs, distance sensors, and pressure insoles to improve spatio-temporal parameter estimation, particularly in detecting complex gait abnormalities such as foot-dragging or freezing episodes. The study emphasized the benefits of multi-sensor integration in reducing drift errors and enhancing real-world gait analysis. In 2023, Salis et al.[49] refined the INDIP system, integrating two plantar pressure insoles, three inertial units, and two distance sensors to comprehensively assess diseased gait in real-world conditions. The system was validated against stereophotogrammetry during both structured laboratory tests and daily-life activity simulations, involving 128 participants from seven cohorts, including healthy individuals and patients with Parkinson's disease, multiple sclerosis, and chronic obstructive pulmonary disease.

Palmerini et al.[50] contributed to the standardization of data collected from wearable devices through the Mobilise-D project. They provided comprehensive guidelines for organizing, integrating, and storing gait data, facilitating the comparison and reproducibility of studies involving IMUs. Their work emphasizes the importance of standardized protocols for enhancing the reliability of digital mobility outcomes across different populations and settings.

Main Results

The application of dedicated IMUs in gait analysis has yielded significant results across various domains. In clinical settings, IMUs have proven invaluable for diagnosing gait abnormalities and monitoring the progression of neurological conditions such as Parkinson's disease and multiple sclerosis. For instance, Bovonsunthonchai et al. [29] demonstrated the effectiveness of IMUs in capturing spatio-temporal

parameters in Parkinson's patients, facilitating early detection of gait impairments and enabling timely interventions.

In rehabilitation, IMUs have been used to track recovery progress and optimize therapeutic strategies. Trojaniello et al.[30] validated the use of IMUs for accurate foot contact detection and spatio-temporal parameter estimation in stroke survivors, highlighting their role in personalized rehabilitation programs. Additionally, IMUs have been employed in sports science to enhance athletic performance by providing detailed feedback on movement mechanics.

Kluge et al.[44] reported that their standardized IMU framework improved the consistency of mobility measurements, enhancing the comparability of data across studies. Mazzà et al.[45] found that IMU-derived gait parameters showed strong correlations with gold-standard laboratory measures, with intraclass correlation coefficients (ICCs) exceeding 0.85 for key metrics like gait speed and stride length. Bertuletti et al.[46] demonstrated that their hybrid MIMU-infrared system provided superior accuracy in step width estimation, with a standard deviation of less than 2 cm compared to traditional IMU setups.

The 2014 study by Salis et al.[47] demonstrated the feasibility of using low-resolution pressure insoles for gait event detection, achieving low root mean square (RMS) errors of 22 ms for initial contacts and 18 ms for final contacts. Their method also showed excellent precision in estimating stance and step duration, with errors below 20 ms and 10 ms, respectively.

Building upon this foundation, the 2019 study [48] further validated the effectiveness of multi-sensor integration, combining IMUs, distance sensors, and pressure insoles to improve the accuracy of spatio-temporal gait assessments. This study highlighted the ability of the INDIP system to detect complex gait abnormalities, such as foot-dragging and freezing episodes, reinforcing its potential for real-world applications.

The latest advancements in 2023 [49] provided a comprehensive validation of the system across diverse clinical populations. The study demonstrated excellent absolute agreement (ICC ≤ 0.95) for all cohorts and digital mobility outcomes during structured tests, with minimal mean absolute errors (cadence ≤ 0.61 steps/min, stride length ≤ 0.02 m, walking speed ≤ 0.02 m/s). Even during real-life simulations, the system maintained high accuracy, with stride length errors ranging from 0.04 to 0.06 m and walking speed errors between 0.03 and 0.05 m/s. These findings underscore the robustness of the INDIP system, confirming its feasibility for long-term gait monitoring in both clinical and free-living conditions.

The versatility of IMUs extends to free-living conditions, where their portability and unobtrusiveness allow for continuous gait monitoring outside of laboratory environments. This capability provides a more comprehensive understanding of an individual's mobility, capturing variations that may not be evident in controlled settings and offering valuable insights for both clinical assessments and daily activity

monitoring.

2.3.3 Limitations of Dedicated IMUs

Despite their numerous advantages, dedicated IMUs present several limitations that can hinder their widespread adoption.

From a practical standpoint, these sensors require periodic calibration to ensure measurement accuracy, often necessitating specialized equipment and technical expertise. Additionally, environmental factors such as temperature variations and electromagnetic interference can introduce noise into the recorded signals, complicating data interpretation. Addressing these challenges frequently demands advanced data processing techniques, including sensor fusion and machine learning, which increase computational complexity and limit their feasibility in certain applications.

Usability is another critical concern, particularly in free-living conditions. The need for customized mounting systems or fixation accessories makes these devices cumbersome and impractical for everyday use. Prolonged wear can also be uncomfortable for users, potentially reducing adherence in long-term monitoring studies.

Cost and accessibility further constrain the adoption of dedicated IMUs. These devices are typically expensive, often costing several hundred euros, which can be prohibitive in resource-limited settings. Moreover, their operation generally depends on proprietary software, requiring manual intervention for data acquisition, sensor calibration, and post-processing. This reliance on specialized tools and expertise introduces an additional layer of complexity, making dedicated IMUs less user-friendly and difficult to integrate into non-specialist settings, where ease of use and automation are essential for scalability.

In summary, while dedicated IMUs remain powerful tools for gait analysis, their cost, usability challenges, and data processing complexity underscore the need for alternative or complementary technologies. The following sections will explore the potential of smartphones as a cost-effective and accessible alternative, addressing some of the challenges associated with dedicated IMUs.

2.3.4 Potential of Smartphones for Gait Analysis

Given the limitations associated with dedicated IMUs—such as high cost, calibration requirements, and usability constraints—smartphones have emerged as a promising alternative for gait analysis. The integration of advanced inertial sensors in modern smartphones has significantly expanded their potential, offering an accessible and cost-effective solution compared to traditional methods relying on standalone IMUs. Equipped with accelerometers, gyroscopes, and magnetometers, smartphones can

capture detailed motion data, enabling the estimation of spatio-temporal gait parameters in both clinical and real-world settings.

One of the key advantages of smartphones over dedicated IMUs is their widespread availability and user familiarity. Unlike traditional IMUs, which often require specialized equipment and setup, smartphones seamlessly integrate into daily life, allowing for unobtrusive and continuous gait monitoring. Their portability makes them ideal for long-term assessments in natural environments, minimizing behavioral modifications that could affect gait patterns.

Moreover, the cost-effectiveness of using smartphones for gait analysis is a major advantage. Since most individuals already own a smartphone, there is no need for additional hardware investments, significantly lowering financial barriers. This democratizes access to gait analysis, making it more scalable and feasible for large-scale studies, as well as for use in resource-limited settings where dedicated IMUs may not be viable. Recent reviews highlight the growing adoption of smartphone-based gait assessment in older adults, emphasizing their utility in both research and clinical applications [51].

2.3.5 Challenges Related to Smartphone Positioning

Smartphone-based gait analysis presents several challenges, primarily due to the variability in device placement, differences in hardware specifications across models, and the influence of user behavior, all of which can introduce inconsistencies in the recorded data. Unlike dedicated biomechanical measurement tools, smartphones are not specifically designed for precise gait analysis, which can impact the accuracy of estimated spatio-temporal parameters. Addressing these challenges necessitates continuous advancements in data processing algorithms to ensure robust and reliable measurements and rigorous validation of smartphone-derived estimates of digital mobility parameters.

A critical aspect influencing the accuracy of smartphone-based gait analysis is device placement. The orientation and placement of a smartphone significantly affect the estimation of spatio-temporal gait parameters. Many biomechanical models used in gait analysis, such as those applied in pedestrian dead reckoning, assume a stable and consistent sensor alignment to estimate step length and gait phases accurately. However, in real-world conditions, smartphones are often carried in diverse locations—such as pockets, hand-held, or inside a bag—leading to significant variations in recorded motion data. These inconsistencies introduce errors in gait parameter estimation, ultimately affecting the reliability of gait assessment algorithms [52].

Despite the growing body of research on smartphone-based motion analysis, most studies have primarily focused on human activity recognition rather than explicitly addressing the impact of smartphone positioning on gait analysis. For

instance, Zhou et al. [53] applied deep learning techniques to classify activities such as walking and stair climbing but did not examine how different smartphone placements influence classification performance. Furthermore, some studies, such as those by Klein [52] and Daniel et al. [54], have attempted to incorporate Smartphone Location Recognition (SLR) into navigation frameworks, showing that position-aware corrections can significantly improve step length estimation. However, these works have been largely confined to controlled laboratory environments, with minimal attention given to the variability and unpredictability of smartphone positioning in real-world scenarios.

Understanding smartphone positioning while walking is crucial for improving the robustness of gait assessment methods, particularly in clinical gait monitoring and rehabilitation. Addressing this challenge enhances the accuracy and reliability of smartphone-based gait analysis, making it more applicable to real-world scenarios and scalable mobility assessments. Despite the challenges related to positioning and measurement accuracy, smartphones offer a cost-effective and widely accessible alternative for gait analysis. The continuous integration of advanced sensor technologies and machine learning algorithms is further refining their precision, solidifying their role as a valuable tool for both clinical research and everyday mobility monitoring.

Chapter 3

Experimental setup

The aim of the work consisted in validating smartphone-derived estimates of digital mobility against a wearable reference system. In the following chapter, the experimental setup of the study, including the smartphone and the reference system, is presented.

3.1 The reference system: INDIP

The INDIP (INertial module with DIstance sensors and Pressure insoles) system is a multi-sensor wearable platform developed by the Università degli Studi di Sassari. Initially conceived as a validation tool in real-world conditions for other gait analysis devices within the Mobilise-D project, INDIP was designed to ensure high reliability at both the hardware and software levels.

3.1.1 INDIP System Setup

The INDIP system consists of inertial sensors, pressure insoles, and distance sensors (Figure 3.1), providing a comprehensive platform for gait analysis. These components can be combined in different configurations to adapt to the specific requirements of various research protocols and experimental conditions. The flexibility of this system allows researchers to select the most appropriate sensor placement based on the parameters of interest and the constraints of each study.

The system includes:

- **MIMUs**, can be placed on various body segments. To ensure proper positioning on the body, the sensors are securely fastened using velcro straps. Those intended for placement on the instep are instead anchored with clips to the shoelaces, ensuring stability during movement.

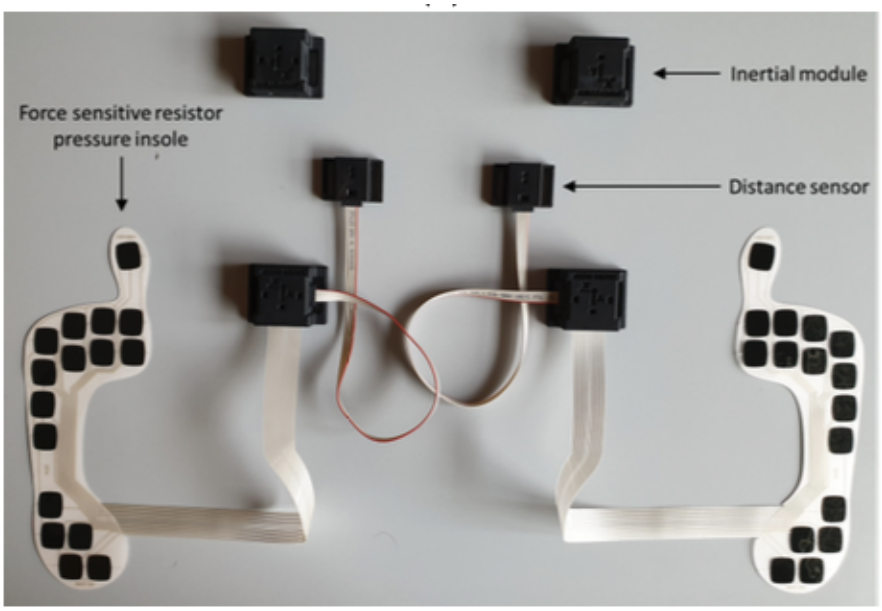


Figure 3.1: Experimental setup of the INDIP system, including two pressure insoles and three MIMUs used as the reference system.

- **Pressure insoles**, designed to be inserted inside the participant's shoes. Each insole contains sixteen pressure sensors, enabling detailed mapping of foot-to-ground contact dynamics, crucial for detecting gait phases and assessing balance.
- **Distance sensors**, which measure inter-feet distance during the gait cycle using infrared technology, providing additional information on step count and width and stride patterns.

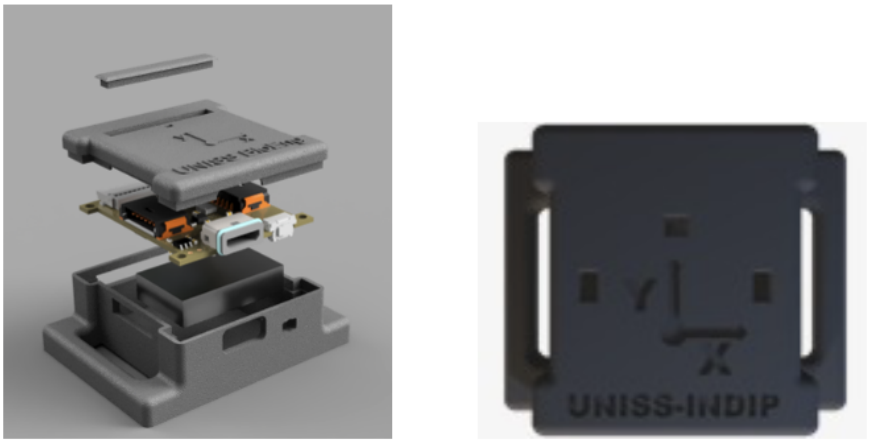
In this study, a configuration of the INDIP system including two pressure insoles and three MIMUs was used as the reference system. Specifically, the setup included:

- **Two pressure insoles**, one for each foot, to assess foot contact timing to ensure better reliability in extracting spatio-temporal parameters.
- **Three MIMUs**, placed as follows:
  - **One at the level of the fifth lumbar vertebra, approximately LB**, position of interest in this study
  - **On the insteps of the right foot and left foot** Sensors were attached to shoelaces via a clip and connected to the pressure insoles via a flexible connector.

This sensor configuration was chosen to ensure reliable gait analysis while minimizing sensor burden on participants, making it suitable for both controlled experimental conditions and potential real-world applications.

3.1.2 INDIP MIMU

Each MIMU consists of a printed circuit board (PCB) that integrates the inertial sensors, transmission modules, and electronic circuitry for data acquisition and storage [46]. The PCB is enclosed within a 3D-printed plastic case (Figure 3.2) and is designed to operate autonomously, featuring an internal battery that provides several hours of continuous use and an onboard memory slot for data storage.



(a) (b)

Figure 3.2: (a) 3-D overview of an INDIP MIMU [49]. (b) Top-view of an INDIP MIMU. The x and y components lie on the sensor's plane, while the z-component points up perpendicularly.

Each MIMU integrates two sensors from STMicroelectronics. The first combines a 3D accelerometer and a 3D gyroscope with selectable full-scale ranges, offering low-power operation and high performance. The second is a 3D digital magnetometer with a ±50 gauss dynamic range, designed for ultralow-power consumption. Additionally, they exhibit low zero-measurement offset and minimal noise, optimizing data accuracy (Table 3.1).

Prior to each recording session, MIMUs must be connected individually via USB to a computer to synchronize their internal timestamps with the current date and time. This process, managed through a custom graphical user interface (GUI) developed in MATLAB, is crucial to ensuring proper alignment between data streams from different sensors. Once synchronization is completed, the MIMUs are attached to the participant and wirelessly connected via Bluetooth Low Energy

TRI-AXIAL ACCELEROMETER	
Measurement range	Up to ±16 g selectable FSR
Zero-g offset	±40 mg
Rate noise density	1.8 - 3.0 mg (Root Mean Squared (RMS))
Output data rate	1.6 to 6664 Hz
TRI-AXIAL GYROSCOPE	
Measurement range	Up to ±2000 dps selectable FSR
Zero-rate offset	±1 dps
RMS noise	0.075 dps
Output data rate	1.6 to 6664 Hz
TRI-AXIAL MAGNETOMETER	
Measurement range	±50 G
Zero-G offset	dynamically cancelled
Rate noise density	3 mG (RMS)
Output data rate	10 to 100 Hz

Table 3.1: Specifications of the sensors of the INDIP MIMU [55]

(BLE) to a custom application, which allows the experimenter to trigger the acquisition process. During recording, participants can move freely. Recorded data can be either streamed in real-time (if the distance between the device and the computer is within the Bluetooth range) or logged to the device memory. At the end of the session, MIMUs are reconnected via BLE to stop the acquisition and then linked to a computer via USB for data transfer using the MATLAB GUI. The device allows to record data at 100Hz or 200Hz, in this study data were recorded at 100Hz.

3.1.3 INDIP pressure insoles

The pressure insole consists of 16 force-sensing resistors embedded in a thin, flexible plastic substrate shaped like a shoe insole [55]. These resistors are strategically positioned along the insole: 9 in the forefoot, 2 in the midfoot, and 5 in the rearfoot (Figure 3.3). Unlike standalone sensors, the pressure insoles do not contain batteries, as they are wired directly to the corresponding foot MIMU for power

and data transmission (Figure **??**). The output data rate of the pressure insoles ranges at 100 or 200 Hz. Pressure insoles are available in different sizes (EU 36-37 40-41 42-43) to comply with the participant's shoe size.

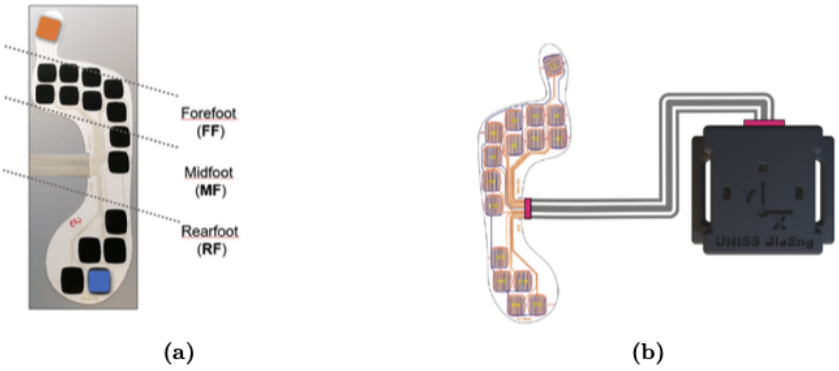


Figure 3.3: (a) Top view of one pressure insole. (b) A pressure insole connected to its corresponding foot MIMU

3.2 The smartphone: Samsung Galaxy A34

The smartphone model selected for this study was the Samsung Galaxy A34. This model was chosen as a representative mid-range smartphone, striking a balance between cost and performance. Given the widespread availability and accessibility of mid-range devices, using such a smartphone allows for a more generalizable approach to gait analysis, ensuring that findings can be applied to a broader population without relying on high-end, expensive models.

However, unlike dedicated IMUs, which have well-documented specifications, consumer smartphones rarely provide detailed technical information about their integrated inertial sensors. Understanding the accuracy, noise levels, and limitations of these sensors is critical for validating their use in gait assessment.

For this reason, a metrological characterization was conducted to determine the Samsung Galaxy A34's inertial sensor specifications, ensuring that its performance could be accurately assessed and compared to a reference system (INDIP). The following sections describe the experimental tests performed, aimed at characterizing the accelerometer, gyroscope, and magnetometer of the smartphone in terms of accuracy, bias, noise, and consistency.

3.2.1 Samsung Galaxy A34: Inertial Sensor Specifications

Due to the lack of publicly available specifications for the Samsung Galaxy A34's inertial sensors, a metrological characterization was carried out to assess their performance. This characterization included a series of tests aimed at evaluating noise, accuracy, bias, and repeatability of the accelerometer, gyroscope, and magnetometer.

Noise Characterization

- **Objective**: Determine the noise level of the accelerometer, gyroscope, and magnetometer.
- **Procedure**:
  - The smartphone was placed on a stable, non-vibrating surface.
  - Data were recorded for two minutes while the device remained static.
  - The noise level was quantified by computing the standard deviation of the recorded signals for each axis. once calculated for each axis, STD values were averaged to provide a unique measure of triaxial standard deviation

Accelerometer Calibration Validation

- **Objective**: Verify the accuracy of the accelerometer under gravitational acceleration.
- **Procedure**:
  - The device was aligned in six standard orientations ($\pm X, \pm Y, \pm Z$) using a rigid support.
  - Each orientation was held for one minute while acceleration data were collected.
  - The measured values were compared to the expected theoretical values (± 9.81 m/s^2 for gravity-aligned axes).

Free-Fall Test for Accelerometer Accuracy

- **Objective**: Validate the accelerometer's performance under pure gravitational conditions.
- **Procedure**:
  - The smartphone was dropped onto a soft surface from a controlled height.
  - The acceleration data were analyzed during the free-fall phase, where the expected total acceleration should be close to zero.

Gyroscope Bias and Stability

- **Objective**: Estimate the gyroscope bias and assess its stability.
- **Procedure**:
  - The device was placed on a stable surface, ensuring no external rotations.
  - Gyroscope data were recorded for two minutes.
  - The mean angular velocity for each axis was computed to assess potential systematic bias.
- **Additional Test**: Gyroscope Rotational Accuracy
  - The device was mounted on a rotating platform (Figure **??**).
  - 100 controlled rotations were executed along each axis.
  - The The integrated angular velocities were compared to the theoretical values to quantify rotational accuracy.

Allan Variance Analysis

- **Objective**: Characterize sensor instability over different time scales [56].
- **Procedure**:
  - the data were collected for two hours in static conditions
  - The collected data were analyzed to identify different noise components, including white noise, bias instability, and random walk.
  - A log-log Allan variance plot was generated to interpret the sensor noise characteristics.

3.2.2 Summary of Findings

All tests were conducted on six Samsung Galaxy A34 smartphones, as these were the same devices used for data collection. Testing all devices allowed for a more comprehensive assessment, ensuring that the results account for device-to-device variability and providing more reliable performance estimates across multiple units.

The *Noise Characterization Test* results, summarized in Table 3.2, provide insights into the sensor stability under static conditions. Standard deviation values were computed for each axis of the accelerometer, gyroscope, and magnetometer across different smartphone placements.

The findings indicate that the accelerometer exhibited standard deviations ranging from 0.0094 to 0.0299, suggesting high stability. The gyroscope showed

Figure 3.4: Rotating platform performing the 100 controlled rotations

a mean standard deviation of 0.0647. The standard deviation for magnetometer ranging from 0.6638 to 5.6417.

For the *Accelerometer Calibration Validation Test* the errors, computed as the

difference between measured and expected values, are summarized in Table 3.3.

The findings reveal that the accelerometer generally aligns well with the expected values, with errors typically remaining below 0.3 m/s^2, except for a few cases where deviations exceeded 0.35 m/s^2 (e.g., +Z orientation for smartphone 1, where the error reached -0.3514 m/s^2). The Y and Z orientations exhibited slightly larger errors, particularly in smartphones 6 and 5.

Under ideal conditions, when an object is in free fall, the accelerometer should measure a total acceleration norm close to 0 m/s^2, as the only acting force is gravity. However, due to sensor limitations, air resistance, and minor external perturbations, the measured norm may deviate from the expected value.

The results of *Free-Fall Test*, summarized in Table 3.4, show that the mean acceleration norms across different smartphone positions range from 0.1324 m/s^2 (smartphone 2) to 0.2458 m/s^2 (smartphone 1). These values, which ideally should be close to zero in a true free-fall scenario, suggest the presence of residual accelerations likely caused by sensor noise, small device movements, or slight imperfections in the free-fall conditions. The mean acceleration norm across all devices is 0.1907 m/s^2, indicating that while the sensors approximate the expected behavior, some residual biases persist.

Ideally, the gyroscope should measure zero angular velocity in a static condition, but in practice, minor biases can occur due to sensor drift, calibration inaccuracies, or inherent noise. These biases can lead to accumulated errors in angular velocity measurements, affecting motion tracking accuracy. The results of *Gyroscope Bias and Stability Test*, summarized in Table 3.5, indicate that the bias values are relatively low across all smartphone positions, with most deviations within ± 0.02 deg/s. The mean bias across all positions is 0.0047 deg/s, suggesting that while individual sensors may exhibit small offsets, the overall bias remains minimal and manageable for gait analysis applications.

In addition to bias evaluation, the *Rotational Accuracy Test* was performed. The recorded angular displacement was then compared to the expected value of 36000°. The results in Table 3.6 show that errors typically ranged between 200° and 350°, indicating a percentage error below 1%.

Figure 3.5 presents the Allan Variance results for different smartphones and sensor axes. The x-axis represents the averaging time τ in seconds on a logarithmic scale, while the y-axis shows the Allan Deviation, which quantifies the sensor's noise characteristics over time.

The observed trends indicate three distinct regions:

- **Short averaging times** ($\tau < 1s$): The Allan Deviation exhibits irregular fluctuations, suggesting a dominant presence of white noise. This high-frequency noise is characteristic of sensor measurement uncertainty and is particularly pronounced in certain axes, such as *Smartphone 1-x* and *Smartphone 3-y*.

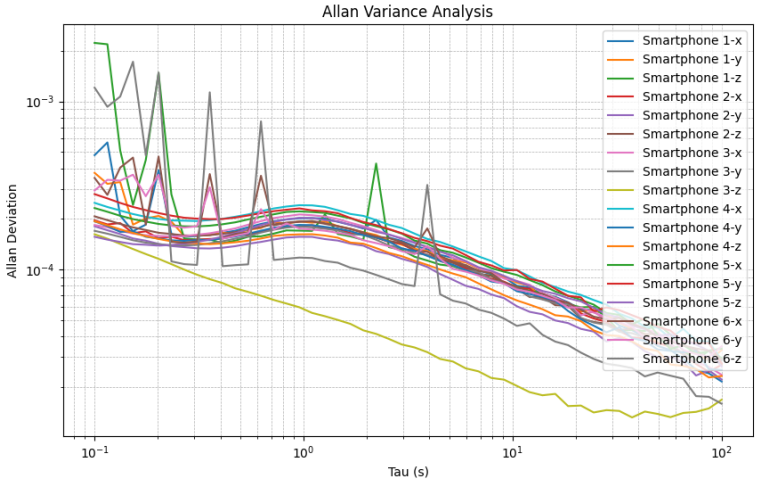


Figure 3.5: Allan Variance analysis for different smartphones and sensor axes (x, y, z).

- **Intermediate region** ($1s < \tau < 10s$): The deviation stabilizes across most axes, indicating a transition from high-frequency noise to bias instability. This suggests that for these timescales, the sensor's performance is less affected by random fluctuations and instead begins to show systematic drift characteristics.
- **Long averaging times** ($\tau > 10s$): Some axes display a continued decrease in Allan Deviation, while others exhibit long-term drift effects. In particular, the *Smartphone 3-z* and *Smartphone 5-z* axes show lower deviation values, implying greater sensor stability in these configurations. Conversely, the oscillations observed in *Smartphone 4-x* and *Smartphone 3-y* suggest increased susceptibility to environmental disturbances or motion artifacts.

Overall, the analysis highlights that sensor placement significantly affects stability, with *Smartphone 3* and *Smartphone 1* exhibiting more consistent noise characteristics over time.

Smartphone	**Sensor**	**Std Dev (Tri-axial)**
1	Accelerometer	0.0094
	Gyroscope	0.0550
	Magnetometer	0.6638
2	Accelerometer	0.0190
	Gyroscope	0.0620
	Magnetometer	3.0358
3	Accelerometer	0.0151
	Gyroscope	0.0540
	Magnetometer	0.9685
4	Accelerometer	0.0148
	Gyroscope	0.0627
	Magnetometer	2.4889
5	Accelerometer	0.0122
	Gyroscope	0.0734
	Magnetometer	5.6417
6	Accelerometer	0.0299
	Gyroscope	0.0812
	Magnetometer	0.8906
Mean	Accelerometer	0.0167
	Gyroscope	0.0647
	Magnetometer	2.2816

Table 3.2: Noise characterization results: triaxial standard deviation values for each sensor and smartphone

Smarthphone	**Orientation**	**Measured (m/s²)**	**Error (m/s²)**
1	-X / +X	9.7061 / -9.9754	-0.1039 / -0.1654
	-Y / +Y	9.7840 / -9.9602	-0.0260 / -0.1502
	-Z / +Z	9.6874 / -10.1615	-0.1226 / -0.3515
2	-X / +X	9.8191 / -9.8700	0.0091 / -0.0600
	-Y / +Y	9.8031 / -9.9578	-0.0069 / -0.1478
	-Z / +Z	10.0612 / -9.8227	0.2512 / -0.0127
3	-X / +X	9.7562 / -9.8434	-0.0538 / -0.0334
	-Y / +Y	9.6692 / -9.9078	-0.1408 / -0.0978
	-Z / +Z	10.0870 / -9.6687	0.2770 / 0.1413
4	-X / +X	9.8658 / -9.8897	0.0558 / -0.0797
	-Y / +Y	9.8447 / -9.9110	0.0347 / -0.1010
	-Z / +Z	10.0035 / -9.8484	0.1935 / -0.0384
5	-X / +X	9.8243 / -9.9155	0.0143 / -0.1055
	-Y / +Y	9.7095 / -9.8616	-0.1005 / -0.0516
	-Z / +Z	10.0731 / -9.6683	0.2631 / 0.1417
6	-X / +X	9.8418 / -9.8039	0.0318 / 0.0061
	-Y / +Y	9.6821 / -9.9573	-0.1279 / -0.1473
	-Z / +Z	9.9703 / -9.8105	0.1603 / -0.0005
Mean Error	-X / +X	-	**-0.0074 / -0.0733**
	-Y / +Y	-	**-0.0616 / -0.1158**
	-Z / +Z	-	**0.1360 / -0.0200**

Table 3.3: Accelerometer Calibration Validation: Measured Acceleration and Errors

Smartphone	**Mean Acceleration Norm (m/s²)**
1	0.2458
2	0.1324
3	0.1983
4	0.1675
5	0.2147
6	0.1856
Mean	**0.1907**

Table 3.4: Free-Fall Test for Accelerometer Accuracy: Measured Acceleration Norm

Smartphone	**Bias x (deg/s)**	**Bias y (deg/s)**	**Bias z (deg/s)**
1	0.0184	0.0167	0.0198
2	0.0159	-0.0174	-0.0165
3	0.0166	-0.0001	0.0160
4	-0.0175	-0.0180	0.0169
5	0.0170	0.0197	-0.0165
6	-0.0185	0.0157	0.0157
Mean Bias	**0.0047**	**0.0028**	**0.0022**

Table 3.5: Gyroscope Bias at Rest: Measured Bias for Each Smartphone

Smartphone	**Error x (deg)**	**Error y (deg)**	**Error z (deg)**
1	243.5	310.2	287.9
2	298.4	275.6	320.1
3	259.7	295.3	305.8
4	276.4	318.9	289.5
5	301.2	250.3	322.8
6	287.9	309.4	270.6
Mean Error	**277.9**	**293.3**	**299.4**

Table 3.6: Gyroscope Rotational Accuracy: Measured Errors in Angular Displacement

Chapter 4

Data collection

The process of data collection is a fundamental step in any study involving gait analysis, as the quality and reliability of the dataset directly influence the validity of the subsequent analyses and model performance. This chapter describes the procedures followed to acquire inertial data from smartphones and the INDIP system in controlled laboratory settings as well as in free-living conditions. The acquisition protocols have been carefully designed to ensure consistency, reproducibility, and a comprehensive representation of real-world walking conditions.

4.1 Smartphone data collection

To record inertial data from smartphones, a custom research-grade application developed by the University of Sheffield, *Mobin*, was used. This application was specifically designed to enable continuous and reliable data acquisition from the smartphones' onboard sensors, ensuring the collection of high-quality signals for subsequent analysis.

Mobin was configured to record data from the accelerometer, gyroscope, and magnetometer at the highest possible sampling frequency supported by the device. The data were logged locally on the smartphone throughout the acquisition session and, once the recording was stopped, the application allowed users to export the recorded data.

The exported data were stored in *CSV format*, with a separate file generated for each sensor. Each file contained four columns:

- **Timestamp** (encoded in milliseconds),
- **X-axis values**,
- **Y-axis values**,

- **Z-axis values**.

This structured format facilitated efficient data handling and preprocessing for subsequent analysis.

Figure 4.1 illustrates the home screen of the *Mobin* application.

4.2 Experimental Protocol

To increase the amount and the heterogeneity of data, two distinct data acquisition sessions were conducted at Politecnico di Torino and at the University of Sheffield. Specifically, data from 15 healthy subjects were collected in Torino (Table 4.2), while an additional 25 healthy subjects were recruited in Sheffield (Table 4.1). Data collection took place both in controlled laboratory settings—*in-Lab*—and in uncontrolled, real-world environments—*free-living*. All participants provided written informed consent for the use of their recorded data. Additionally, the data collection process conducted in Sheffield received ethical approval from the Ethics Committee of the University of Sheffield.

In addition to inertial data, demographic information and specific measurements were collected from participants, including the height of the LB-MIMU and smartphone from the ground. These measurements were required for the correct application of the Mobilise-D pipeline, ensuring accurate processing and validation of the extracted gait parameters.

This chapter details the experimental setup, data acquisition protocols, and preprocessing steps necessary to prepare the collected signals for further analysis

4.3 Acquisitions in Torino

The acquisitions done in Torino were conducted in a controlled enviroment, the experiments took place within the Department of Electronics and Telecommunications (DET) at Politecnico di Torino.

4.3.1 Data acquisition protocol

Participants were equipped with the INDIP kit configuration previously described in Section 3.1.1. Additionally, a smartphone was securely positioned on the lower back using an elastic belt to ensure stability during walking. The smartphone was placed in a horizontal orientation, with its sensor axes aligned as follows:

- The **positive Z-axis** aligned with the anteroposterior direction.
- The **positive X-axis** aligned with the inferior-superior direction.

Subject	Gender	Age	Height (cm)	Shoe Size (EU)
1	M	36	185	44
2	M	26	192	46
3	F	27	158	38
4	M	28	180	44
5	F	27	171	40
6	M	57	174	42
7	M	29	182	45
8	M	38	178	43
9	F	31	165	37
10	F	27	166	38
11	M	42	170	42
12	M	26	169	40
13	M	29	187	45
14	M	34	175	43
15	M	30	176	43
16	M	31	186	44
17	M	40	175	43
18	M	30	175	42
19	F	33	152	39
20	F	29	164	38
21	M	24	175	44
22	M	30	184	43
23	M	37	174	42
24	F	28	173	37
25	F	22	166	39
Summary	M (68%)	31.64±7.22	174.08±9.31	41.50±3.00

Table 4.1: Participant characteristics for the Sheffield data acquisition

- The **positive Y-axis** aligned with the mediolateral direction.

The described equipment setup, along with an illustration of the reference coordinate system used in Android smartphones, is presented in Figure 4.2.

The in-lab data acquisition protocol consisted of seven structured tests, each conducted under the supervision of one or two operators. Each test included a pre-defined number of repetitions, referred to as *Trials*. Below, a brief overview of each test is provided.

- **Spot check test (Test 1):** This test aimed to assess the quality of sensor performance. Data collected during this phase were analyzed using the INDIP Data Quality Check tool to detect any artifacts or inconsistencies.

Subject	Gender	Age	Height (cm)	Shoe Size (EU)
1	M	28	185	45
2	M	25	179	46
3	M	25	190	46
4	F	29	170	37
5	F	27	173	39
6	M	25	176	43
7	F	23	161	38
8	M	24	176	43
9	M	24	175	44
10	F	25	165	39
11	F	24	175	40
12	M	25	179	42
13	M	34	180	43
14	F	24	162	38
15	M	28	180	44
Summary	M (60%)	26.00±2.83	175.06±8.03	41.08±3.05

Table 4.2: Participant characteristics for the Torino data acquisition

- *IMU static test:* A static acquisition was performed on each INDIP MIMUs that was utilized during the acquisition. The devices were placed on a flat surface for at least 60 seconds while ensuring no movement or disturbances affected the sensors.
- *Pressure insoles static test:* This test verified the correct functionality of the pressure insoles. An operator sequentially applied pressure to each sensing unit, independently for both the right and left insoles.

The data recorded during this test were used for quality control. For the MIMUs, the system evaluated whether the inertial signals had an expected mean value (approximately zero), an accelerometer norm within the expected range, and an appropriate standard deviation. Additionally, the system checked whether the full-scale range was correct or if any sensors had ceased recording during the acquisition. To ensure optimal performance during the IMU static test, recording should not commence immediately after powering on the MIMUs. This precaution is necessary because the sensors require a stabilization period to reach their steady-state operating conditions. Failure to allow sufficient time for stabilization often resulted in the spot check failing, with the system reporting a "high standard deviation" condition. Similarly, for the pressure insoles, the system assessed potential performance degradation or recording failures. If any issues were identified, the faulty sensors were

replaced, and the test was repeated before proceeding.

- **Standing test (Test 2):** This test involved a brief static acquisition, during which the participant wore the complete device setup, including MIMUs, pressure insoles, and the smartphone. The participant was instructed to stand still for at least 10 seconds. This phase was essential for estimating the angle between vertical axis of the device and the gravity axis, enabling the computation of the rotation matrix required to align the inertial recordings during subsequent dynamic acquisitions (Figure 4.4).
- **Data Personalization (Test 3):** This phase ensured the proper placement of the pressure insoles inside the participant's shoes and identified any malfunctioning sensors that required recalibration or substitution. The participant was instructed to:
  - Stand still for at least 10 seconds.
  - Lift the left foot for at least 5 seconds (single right-leg support).
  - Return to a double support stance for 5 seconds.
  - Lift the right foot for 5 seconds (single left-leg support).
  - Return to a double support stance for 5 seconds.
  - Walk at a comfortable pace along a 12-meter straight path(Figure 4.3).

  Signals recorded during this test are illustrated in Figure 4.5.
- **Slow straight walking (Test 4):** The participant walked along a 12-meter straight path at a slow pace. This test was repeated three times (Figure 4.6).
- **Normal straight walking (Test 5):** The participant walked along the same 12-meter path at a self-selected comfortable speed, again repeating the test three times (Figure 4.7).
- **Fast straight walking (Test 6):** The participant walked the same path at an increased pace. This test was also repeated three times (Figure 4.8).
- **Round walking (Test 7):** The participant walked four full laps around an approximately 24-meter oval (Figure 4.3) path at a comfortable speed. This test was repeated three times and was included in the protocol to introduce non-linear walking trajectories into the dataset (Figure 4.9).

Conducting each test three times ensured redundancy, allowing for data recovery in case of corrupted trials. Additionally, incorporating different walking speeds and non-linear trajectories increased dataset variability, enhancing the generalization capability of machine learning models. However, this also introduced the risk of underfitting, as the model was required to learn a broader range of walking patterns.

4.3.2 Data preparation

Upon completion of all tests, the data recorded from all devices were manually downloaded and systematically organized into structured directories to facilitate preprocessing. Before standardizing each participant's files into a MATLAB structure (*data.mat*), preliminary data processing was performed.

The signals recorded by the INDIP MIMUs were inherently structured per trial, allowing for a direct mapping of each test repetition. Conversely, due to the inability to pause and resume recordings remotely, the smartphone data were acquired as a single continuous recording. Consequently, a segmentation method was implemented to identify and extract individual trials from the smartphone recordings.

This segmentation was achieved by leveraging cross-correlation analysis between the norm of the gyroscope signal from the LB-MIMU and the norm of the gyroscope signal from the entire smartphone recording. The time instants corresponding to the highest cross-correlation values were identified, and the time delay between the two signals was estimated and applied to the smartphone data to synchronize it with the MIMU reference signal. Subsequently, the smartphone recording was segmented to isolate the signal corresponding to each trial, as illustrated in Figure 4.10.

4.4 Acquisitions in Sheffield

The acquisitions done in Sheffield were conducted in both controlled and free-living confditions. In these acquisitions, participants were equipped as previously described in Section 4.3.1 and illustrated in Figure 4.2.

To ensure comprehensive inertial data collection for the smartphone location recognition task, participants carried six smartphones simultaneously, each assigned to a predefined position, as shown in Figure 4.11. This setup allowed for the development and evaluation of a model capable of recognizing smartphone placement during gait. Among these, the LB position was specifically included, as the Mobilise-D algorithm pipeline is designed for inertial data from this location. The remaining five positions were selected based on common real-world smartphone placements, as detailed in Section 3.2.2, ensuring a representative dataset for the recognition task.

4.4.1 Data acquisition protocol

Immediately after performing the spot check test, an additional step was introduced to ensure a common temporal reference for the signals recorded by the six smartphones. Before being worn by the participant, all six smartphones were placed on a support mounted on a turntable and rotated simultaneously. This procedure

generated an initial gyroscope signal with a consistent angular velocity pattern across all devices, which was later used for temporal realignment.

This step was necessary due to a key difference between the experimental setups in Torino and Sheffield. In the Torino acquisitions, only one smartphone was placed on the lower back alongside an INDIP MIMU, ensuring that both recorded the same movement as they shared the same anatomical reference point. However, in the Sheffield configuration, where six smartphones were placed in different locations, each device recorded distinct motion patterns. The turntable-based synchronization approach was therefore implemented to provide a common reference point, facilitating accurate temporal realignment across all smartphones.

Once the previously described tests were completed in a structured environment (Section 4.3.1), participants were released into a free-living context for a duration of 2.5 hours. Given the nature of free-living conditions, participants were not required to walk continuously for the entire period; instead, they were granted full autonomy to carry out their daily activities. During this time, they were allowed to remove the jacket and shoulder bag, as well as temporarily take off all smartphones and place them on a surface, except for the smartphone positioned on the lower back and the INDIP system.

Participants were instructed that whenever they resumed walking, they should reattach all smartphones to their designated positions. To facilitate the correct reattachment, each smartphone was labeled according to its specific placement, and the background of each device was customized accordingly, as illustrated in Figure 4.12.

Although participants were given complete freedom in their activities, they were encouraged to walk as much as possible throughout the session to maximize the amount of data collected for the development of the classifier.

4.4.2 Data preparation

The process for generating the MATLAB structure *data.mat* [50] followed the same steps outlined in Section 4.3.2, with additional procedures required due to the presence of multiple smartphones. Since six smartphones were used simultaneously, it was necessary to align all recordings to a common temporal reference.

To achieve this, all smartphones were synchronized with the lower back smartphone using cross-correlation. Specifically, the angular velocity pattern recorded during the controlled rotation on the turntable was leveraged as a shared reference signal. This ensured that the initial timestamps of all smartphone recordings were aligned before further processing.

Once the smartphones were temporally synchronized, the next step was to align the LB smartphone with the LB-MIMU. Maintaining continuous synchronization between the LB smartphone and the LB-MIMU was crucial, as the INDIP pipeline

relied on the LB-MIMU signals to segment walking bouts, using its timestamps to define the start and end of gait segments. This alignment ensured that all extracted gait segments were correctly referenced across devices.

Finally, the *data.mat* structure was populated not only with all structured test data—including the entire 2.5-hour free-living session—from both the LB smartphone and the INDIP system, but also with the data from all additional smartphones.

4.4.3 Intended use of the dataset

Data acquired in Turin and Sheffield in a controlled environment resulted in the *In-Lab* dataset. The one acquired in the free-living context in Sheffield resulted in the *Free-living* dataset. The acquired datasets served two primary objectives. First, for the validation of the Mobilise-D algorithm pipeline, the INDIP kit was used as the reference system to extract the temporal values of DMOs through its dedicated algorithm. These reference values were then compared against the DMOs estimated from the inertial data recorded by the smartphone, which serves as the device under test, and processed through the Mobilise-D algorithm pipeline. The second objective was to develop a model for smartphone position recognition during gait, leveraging the multiple smartphone placements included in the dataset. This setup ensured the availability of diverse inertial data necessary to train and evaluate the recognition model, enhancing its robustness across different carrying positions.

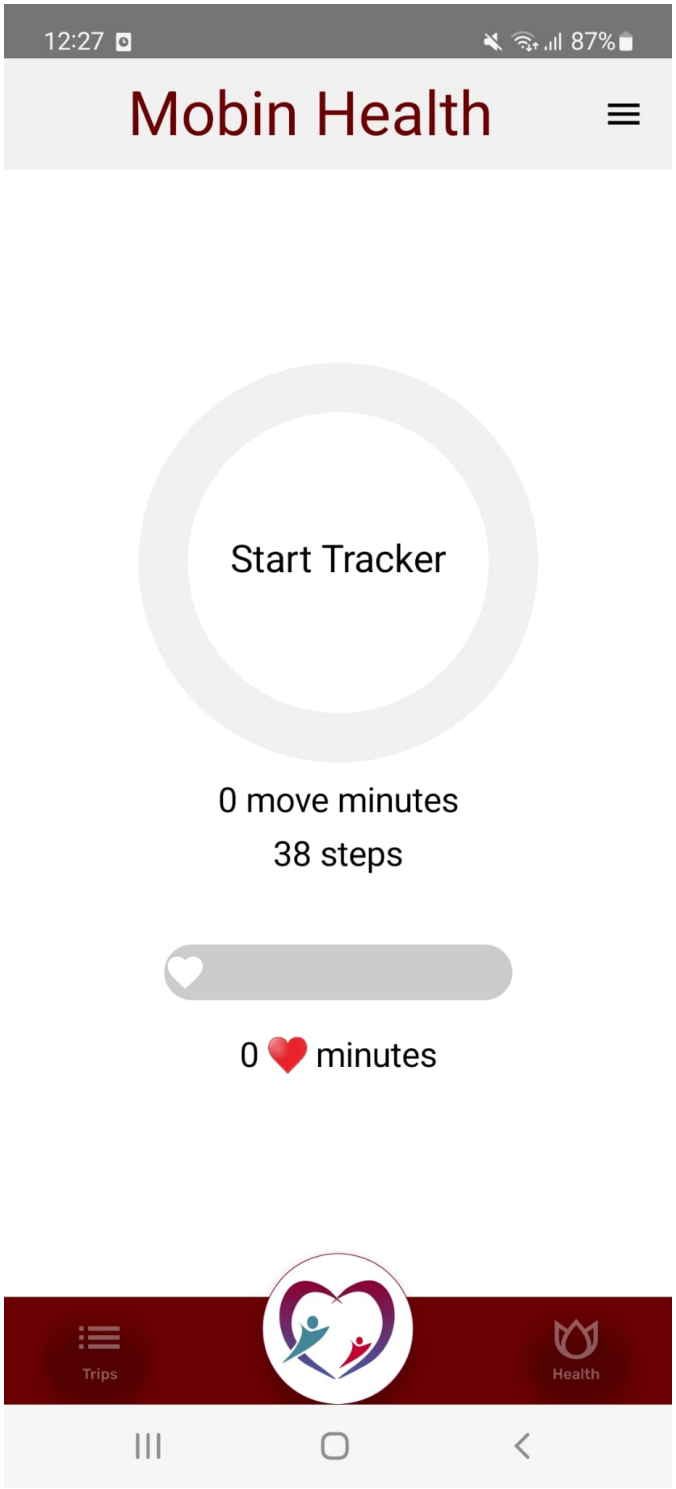


Figure 4.1: Home screen of the *Mobin* application used for smartphone data collection.

Figure 4.2: Standard reference coordinate system in Android smartphones, experimental setup configuration with a focus on foot-mounted MIMUs, the smartphone, and the LB-MIMU.

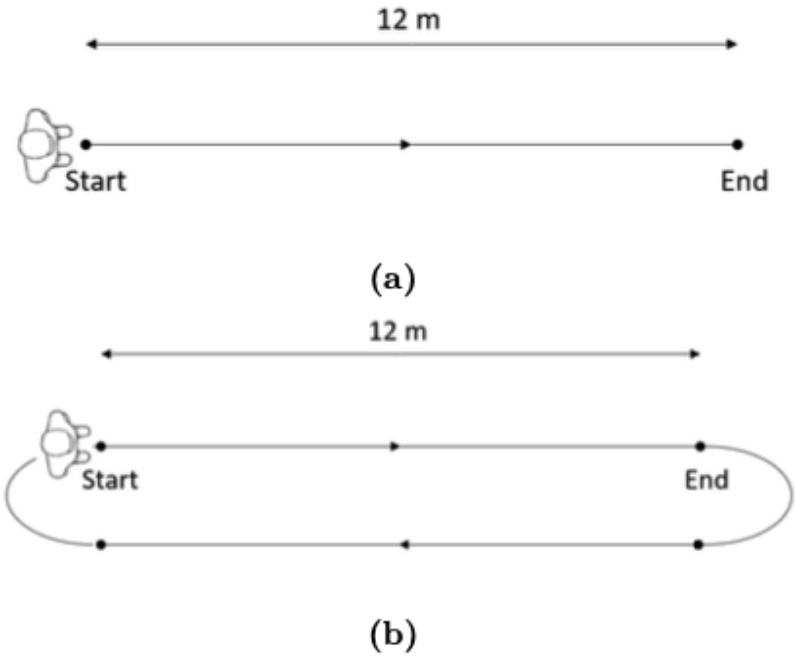


Figure 4.3: (a) Straight path for Test 3, 4, 5 and 6. (b) Ring path for Test 7.

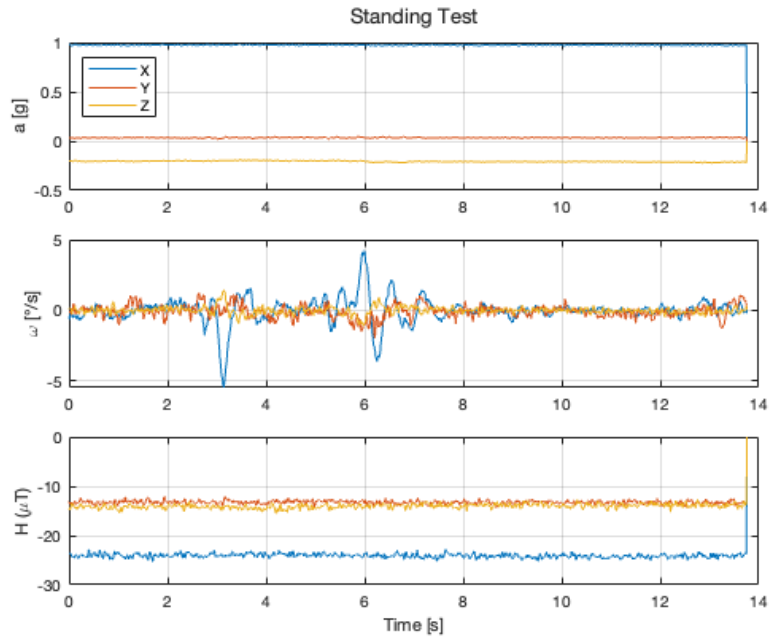


Figure 4.4: Acceleration (a), Angular velocity (ω) and magnetic field (H) recorded from the smartphone positioned on the lower back during the Standing Test.

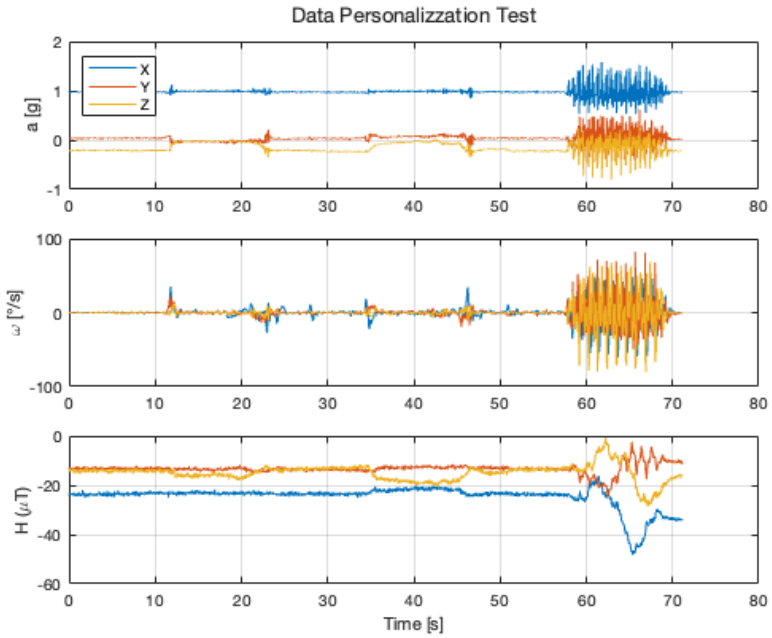


Figure 4.5: Acceleration (a), Angular velocity (ω) and magnetic field (H) recorded from the smartphone positioned on the lower back during the Data Personalization Test.

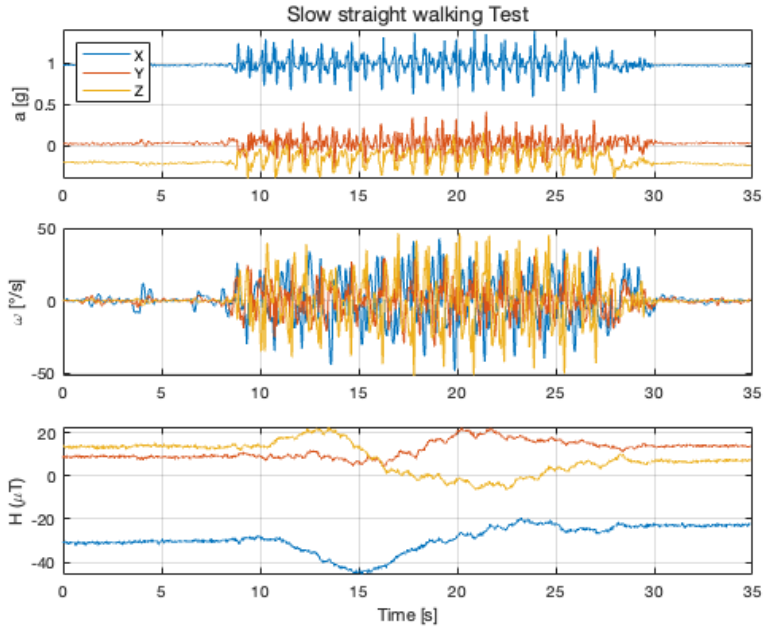


Figure 4.6: Acceleration (a), Angular velocity (ω) and magnetic field (H) recorded from the smartphone positioned on the lower back during the Slow Straight Walking Test.

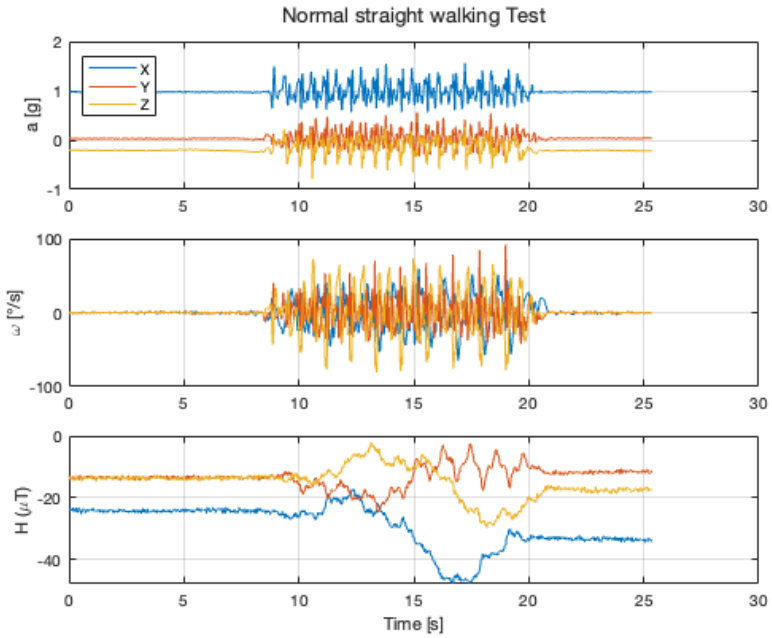


Figure 4.7: Acceleration (a), Angular velocity (ω) and magnetic field (H) recorded from the smartphone positioned on the lower back during the Normal Straight Walking Test.

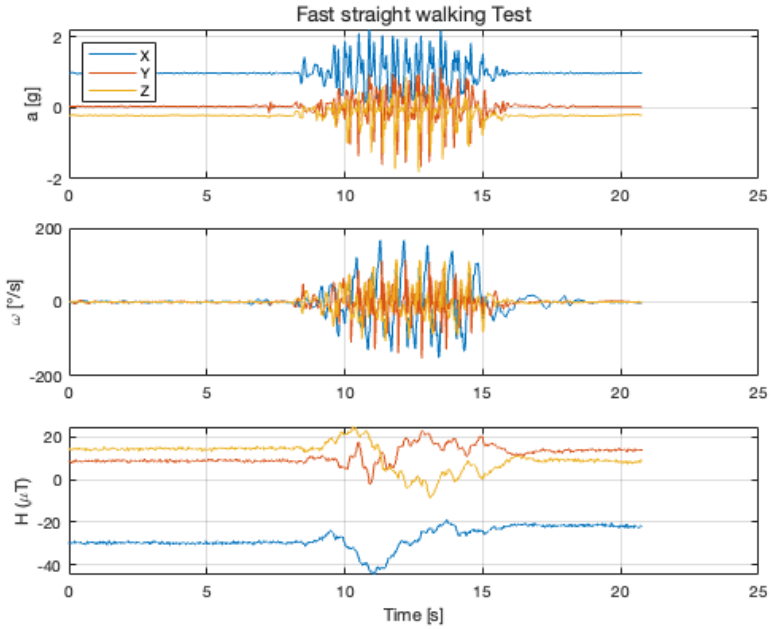


Figure 4.8: Acceleration (a), Angular velocity (ω) and magnetic field (H) recorded from the smartphone positioned on the lower back during the Fast Straight Walking Test.

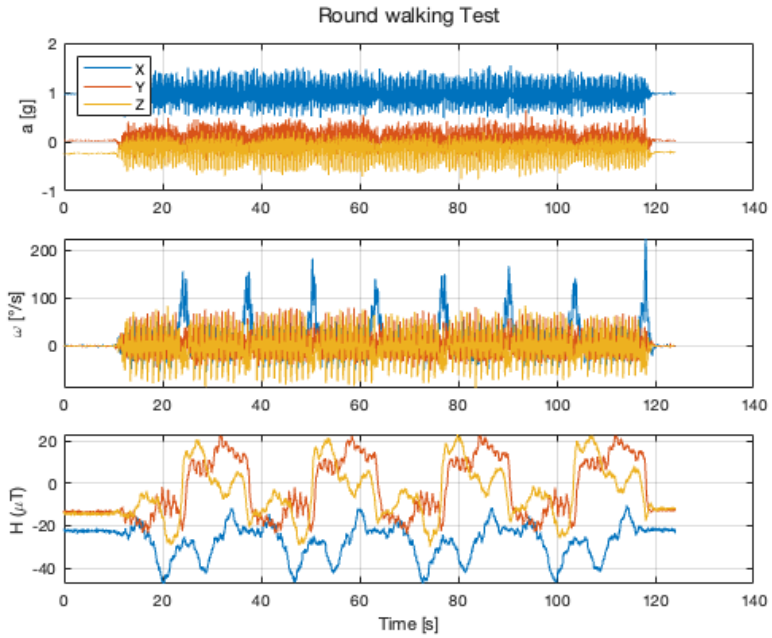


Figure 4.9: Acceleration (a), Angular velocity (ω) and magnetic field (H) recorded from the smartphone positioned on the lower back during the Round Walking Test.

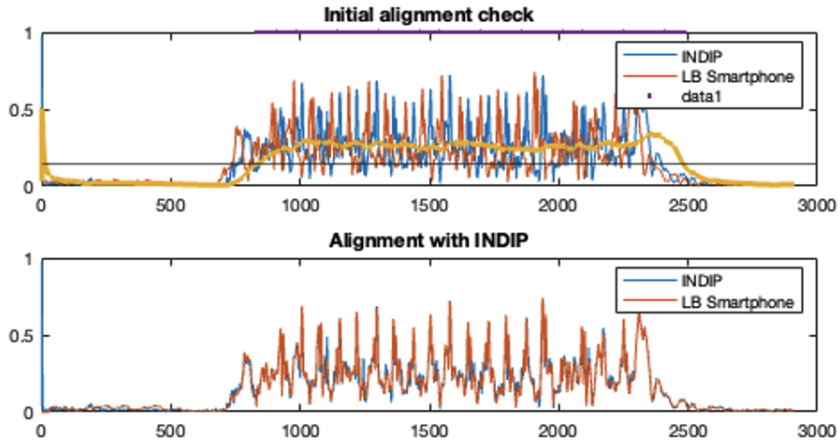


Figure 4.10: Comparison of signal alignment between the LB Smartphone and the INDIP system. The top plot shows the initial alignment check, highlighting discrepancies between the signals. The bottom plot illustrates the improved alignment after synchronization with INDIP.

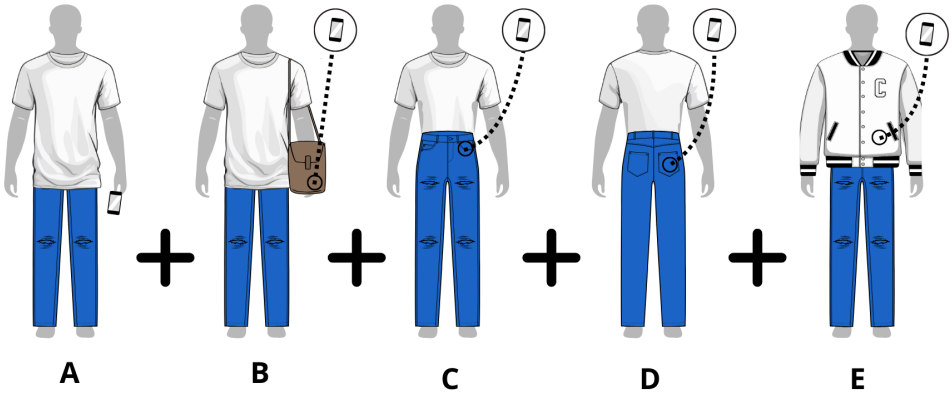


Figure 4.11: Schematic representation of the five additional smartphone positions investigated in the SLR task: (A) Hand-held, (B) Shoulder bag, (C) Front pocket, (D) Back pocket, and (E) Coat pocket.

Figure 4.12: Customized smartphone backgrounds indicating their designated positions to ensure correct reattachment by participants.

Chapter 5

Methods

5.1 Overview

As, stated in previous chapters, this thesis focuses on two main objectives:

- **Validation of the Mobilise-D algorithms and pipeline with smartphone-acquired inertial data.** The study assessed whether the Mobilise-D single algorithms and full pipeline, originally validated for a lower back-mounted dedicated IMU, can be reliably applied to smartphone inertial data. Key gait parameters were evaluated and compared with reference values derived from the INDIP system.

- **Development and validation of a machine learning framework for smartphone location recognition.** The study assessed the feasibility of recognizing the smartphone location during walking. Five machine learning models were supervisedly trained using labeled inertial data recorded from multiple smartphones simultaneously and their performance compared to assess the optimal architecture and features set for the task. This process involved constructing a labeled dataset from multiple smartphone placements, extracting relevant features from accelerometer and gyroscope signals, and training a classification model to distinguish between predefined carrying positions.

5.2 MobGap validation with smartphone data

The Mobilise-D algorithm pipeline is an analytical framework developed to assess real-world gait and estimate Digital Mobility Outcome (DMOs) from wearable inertial sensors. Originating from the Mobilise-D project, a large-scale European

initiative, it provides a standardized and validated method for extracting and analyzing gait parameters from a single IMU mounted on the lower-back.

5.2.1 Development and validation of the Mobilise-D algorithm pipeline

The development and validation of the Mobilise-D algorithm pipeline was performed in the context of the Technical Validation Study (TVS) of the Mobilise-D project [45]. The goal of the TVS was to demonstrate the technical validity of an algorithm pipeline for the assessmnt of DMOs in real-world conditions. Such DMOs, including gait sequences, initial contacts, cadence, stride length and walking speed, were dependent one on each other (e.g., stride length can be estimated after that initial contacts are detected). Therefore, the algorithm was devised as a *pipeline* composed of various *blocks*, each one associated to many algorithms [1]. To identify the best-working algorithm for each block, a block-by-block validation of the algorithm pipeline was performed [11], resulting in a final implementation referred to as the Mobilise-D algorithm pipeline [9]. Recently, an open-source Python implementation of the pipeline called MobGap has been validated and released on GitHub [10].

5.2.2 Technical Validation Study dataset

During the TVS, the pipeline, as well as the single algorithms, have been previously validated with data of healthy individuals and patients with mobility impairments, including both healthy individuals and those with mobility impairments, ensuring its applicability for both clinical and general use [11][9]. One of the key strengths of the Mobilise-D pipeline is its validation across multiple clinical conditions that affect gait and mobility. Specifically, it has been validated in the context of the Technical Validation Study (TVS)[45], with data of individuals with:

- **Parkinson's disease**: Reduced gait speed, shorter step length, increased gait variability, and freezing of gait episodes.
- **Multiple sclerosis**: Gait asymmetry, impaired coordination, reduced cadence, and difficulty in maintaining a stable walking pattern.
- **Chronic obstructive pulmonary disease**: Reduced endurance, altered step dynamics, and increased gait variability due to respiratory limitations.

[1] In the original implementation of the pipeline, most algorithms were coded in Matlab, other in Python, only a few in R.

- **Proximal femoral fracture**: Impaired weight-bearing, reduced stride length, and compensatory movement patterns.
- **Congestive heart failure**: Affected walking speed and cadence, contributing to slower and more effortful gait patterns.
- **Older adults with frailty**: Decreased step regularity, reduced stride length, and increased risk of falls [9].

A refined version of the dataset has been recently released on the open-source platform Zenodo [57]

Beyond clinical populations, the Mobilise-D algorithm pipeline is also suitable for healthy individuals, enabling objective gait monitoring in everyday settings. Unlike traditional gait analysis methods, which are often restricted to laboratory environments, Mobilise-D allows for continuous gait assessment also in free-living conditions, offering ecologically valid insights into mobility [9].

5.2.3 Blocks of the Mobilise-D algorithm pipeline

The pipeline comprises multiple algorithmic components, each responsible for processing specific aspects of gait data (Figure 5.1).

The Mobilise-D algorithm pipeline is based on the following key blocks [11]:

- **Gait Sequence Detection (GSD)**: Identifies periods of walking within continuous sensor recordings, distinguishing gait sequences from non-walking activities. This step is fundamental as it defines the segments where gait analysis is performed. Detected gait sequences include straight as well as curvilinear walking, in both inclined and flat surfaces.
- **Initial Contact Detection (ICD)**: Determines the precise moments when a foot makes contact with the ground, a key spatio-temporal parameter. The refined gait sequences are obtained by detecting the start and end of walking phases based on these contacts. Additionally, this spatio-temporal parameter is essential to estimate step and stride according to what previously defined
- **Cadence Estimation (CAD)**: Measures the number of steps per minute. It is computed on a per second basis within each detected gait sequence, providing a key measure of walking rhythm and intensity.
- **Stride Length Estimation (SL)**: Computes stride length by measuring acceleration-derived vertical displacement and assuming a biomechanical model of the lower body, providing a stride length estimate per second of gait.

In addition to these fundamental components, the pipeline includes other modules that further refine the gait analysis and provide complementary parameters:

- *Turning Detection*: Identifies turns during walking and provides details on the start, end, and degree of turning within the detected gait sequence.
- *Left-Right Foot Classification*: Assigns detected initial contacts to the left or right foot, enabling gait asymmetry analyses.
- *Walking Speed Calculation*: Computes the walking speed per second using cadence and stride length estimates.
- *Stride-Level Interpolation*: Ensures all calculated parameters are interpolated to a fixed time step to maintain temporal consistency across estimates.
- *Stride Selection*: Filters out outlier stride estimates that fall outside physiologically plausible ranges, improving the robustness of extracted gait parameters.
- *Walking Bout Assembly*: Segments and structures gait data into meaningful walking bouts, defining their start and end based on stride sequences.

Each algorithm in the pipeline has been optimized and validated against gold-standard reference systems, such as INDIP, in the configuration as previously described in Section 5.2.2 to ensure high-fidelity ground truth data [9].

5.2.4 Algorithms of the Mobilise-D algorithm pipeline

In this section, the algorithms recommended for correctly applying the Mobilise-D pipeline to healthy individuals are analyzed and described. Based on the findings from the technical validation study [9] [11], specific algorithmic choices ensure optimal performance when processing gait data from a population of healthy subjects, which is the target population for this study.

For this purpose, the most suitable algorithms for each of the core pipeline blocks have been identified. According to the validation study, the recommended algorithms for healthy individuals correspond to those within the pipeline configuration labeled as P1 in Figure 5.1. These algorithms demonstrated the best accuracy and reliability in detecting and estimating key gait parameters under free-living conditions.

In this study, the Python implementation of the Mobilise-D algorithm pipeline was employed, which is available through the *MobGap* package. The version employed in this study is *MobGap 0.10.0*, which includes a validated and structured set of methods for gait sequence detection, initial contacts detection, cadence estimation, stride length estimation and walking speed estimation. The use of this

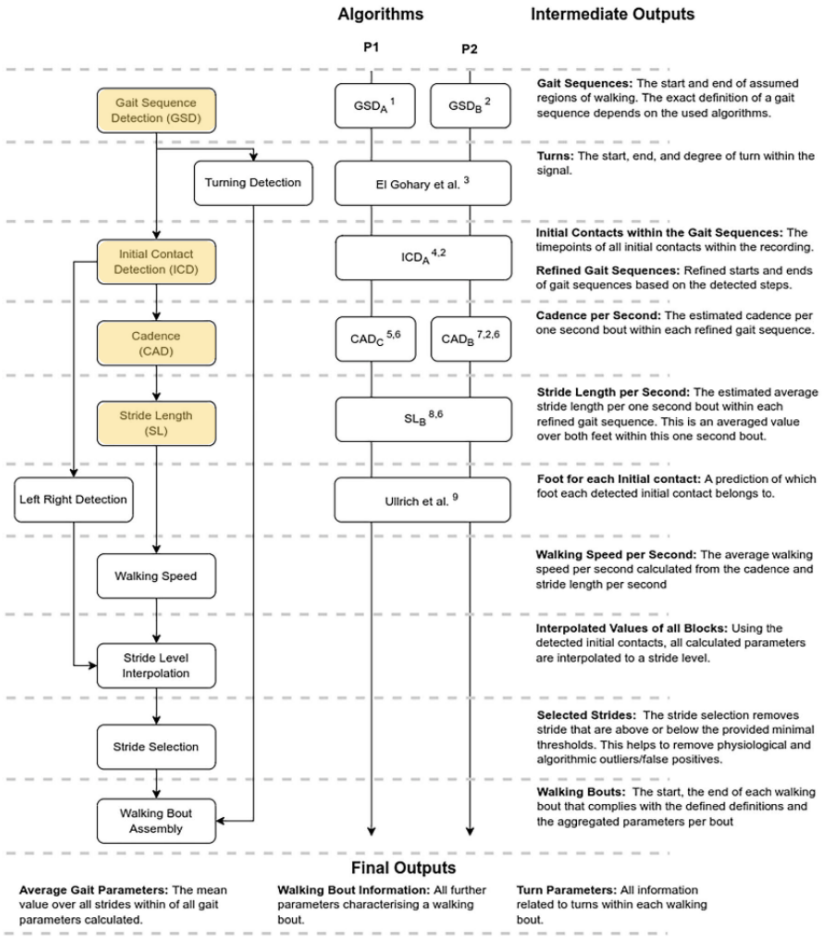


Figure 5.1: Schematic representation of the Mobilise-D pipeline, outlining its main algorithmic components and their respective outputs. The main validated blocks are highlighted.Adapted from Kirk et al. [9]

package ensures reproducibility and compliance with the standardized pipeline applied in previous studies.

The following subsections provide a detailed explanation of each selected algorithm and its role within the pipeline.

Gait Sequence Detection

The detection of gait sequences is a fundamental step in the Mobilise-D pipeline, ensuring that only valid walking periods are analyzed while filtering out non-walking activities such as standing, sitting, or external perturbations. The method follows

the approach proposed by Iluz et al. (2014) [58], leveraging the characteristic frequency content of human gait, typically within the 0.5–3.0 Hz range. A band-pass filter is applied to both vertical and anteroposterior acceleration components to extract gait-related motion while attenuating noise.

A sliding window approach is then used to detect periodic gait cycles:

- A 5-second moving window is convolved with a 2 Hz sinusoidal reference signal, representing the periodicity of human gait.
- Local maxima in the resulting signal correspond to detected gait cycles, segmenting walking periods.
- Windows with a step frequency outside the physiologically plausible range of 2–15 steps per window are discarded.

Building on this approach, the implementation in the `GsdIluz` class within the MobGap package introduces refinements to improve computational efficiency and robustness [59]. The main modifications include reordering the processing steps to optimize computational cost, replacing the custom peak detection method with the `find_peaks` function from the Python library `SciPy`, and unifying thresholding across sensor axes. Additionally, all parameters are converted to SI units, and adjustments have been made to correct threshold applications to prevent erroneous peak suppression.

Initial Contact Detection

The detection of initial contact (IC) events is critical for temporal gait analysis, as it enables the computation of parameters such as stride time and cadence. The method follows the approach of McCamley et al. (2012) [60], later refined by Paraschiv-Ionescu et al. (2020) [61], and operates as follows:

- The accelerometer signal, recorded at 100 Hz, is downsampled to 40 Hz for computational efficiency.
- A band-pass filter (0.15–3.14 Hz) isolates gait-related components while suppressing noise.
- The filtered signal undergoes cumulative trapezoidal integration, followed by a Continuous Wavelet Transform using the *Ricker wavelet*.
- Zero crossings are detected, and negative peaks occurring between crossings are classified as IC events.

In this study, the MobGap implementation of IC detection, through the *IcdIonescu* class, retains the core methodology while incorporating minor refinements. These modifications include an improved downsampling strategy and a refined peak detection approach, which enhances computational efficiency while preserving consistency with the original MATLAB implementation [62].

Cadence Estimation

Cadence estimation quantifies step frequency over time, providing insights into gait stability and rhythm. The method extracts step times from detected ICs and follows these steps:

- Step times are calculated from IC events. The last step time is replicated to ensure each IC has an associated step time.
- A Hampel filter smooths the step time series, removing outliers.
- Step times per second are computed by averaging step times within each second. If no ICs occur in a second, missing values are linearly interpolated unless the gap exceeds a predefined threshold, in which case NaNs are assigned.
- A second smoothing step is applied before cadence is computed as the inverse of step time per second.

The MobGap implementation modifies the original MATLAB method by postponing cadence computation until all preprocessing steps are complete. This avoids distortions due to outliers and ensures greater robustness. Additionally, a maximum interpolation gap is introduced to prevent erroneous assumptions about missing data, and linear interpolation is used instead of simple averaging to better preserve temporal structure [63].

Stride Length Estimation

Stride length estimation is essential for characterizing spatial gait parameters and is based on an inverted pendulum model. The method follows the approach by Zijlstra (2003) [26], later refined by Soltani et al. (2021) [64], and proceeds as follows:

- Sensor signals may be aligned using a Madgwick complementary filter (optional).
- A 4th-order Butterworth high-pass filter (0.1 Hz cut-off) removes drift from acceleration signals.

- Vertical acceleration is integrated to obtain vertical speed, followed by a second high-pass filter (1 Hz) to remove drift.

- A second integration yields vertical displacement $d(t)$.

- The total vertical displacement during a step (d_{step}) is computed as:

$$d_{step} = |max(d(t)) - min(d(t))| \tag{5.1}$$

- Stride length is estimated using a biomechanical model:

$$\text{StrideLength} = A \cdot 2 \cdot \sqrt{2 \cdot LBh \cdot d_{\text{step}} - d_{\text{step}}^2} \tag{5.2}$$

  where A is a tuning coefficient optimized through grid search, and LBh represents the approximate sensor height relative to the center of mass.

Stride length estimates undergo further post-processing:

- Step lengths are computed from ICs, and missing values are handled by replicating the last step length.

- A Hampel filter smooths the step length data.

- Step length per second is computed, with missing values linearly interpolated unless the gap exceeds a predefined threshold.

- A final smoothing pass is applied, and stride length per second is obtained by doubling step length.

The MobGap implementation introduces refinements to the original method, including a maximum interpolation gap to avoid masking errors in IC detection and linear interpolation instead of simple averaging to improve robustness against outliers [65].

Additionally, to ensure that the selected algorithms from pipeline P1 (Figure 5.1) were indeed the best suited for this specific dataset, an extensive comparative evaluation was performed. All available algorithms within MobGap (Table 5.1) for each processing block were tested and evaluated based on their accuracy and agreement with the reference system. This analysis aimed to verify whether the predefined P1 configuration remained the optimal choice in this context or if alternative algorithms provided better performance.

Module	Available Algorithms
Gait Sequence Detection (GSD)	GSD_A (Illuz et al., 2014)
	GSD_{A2} (Modified implementation of GSD_A)
	GSD_B (Ionescu et al., 2019)
	GSD_C (Modified implementation of GSD_B)
Cadence Estimation (CAD)	CAD_A (Lee et al., 2010)
	CAD_B (Ionescu et al., 2011)
	CAD_C (Shin et al., 2011)
Initial Contact Detection (ICD)	ICD_A (Ionescu et al., 2019)
	ICD_D (Shin et al., 2011)
	ICD_E (Lee et al., 2010)
Stride Length Estimation (SL)	SL_A (Zijlstra et al., 2013)
	SL_B (Modified implementation of SL_A)

Table 5.1: Summary of available algorithms in the MobGap pipeline, categorized by module. Modified implementations are indicated accordingly.

5.2.5 Validation Strategy

To validate the Mobilise-D algorithm pipeline, two complementary approaches were employed: a block-by-block validation and a full-pipeline validation. Each method provides distinct insights into the accuracy and reliability of the system.

The block-by-block validation, inspired by the methodology used by Micò-Amigo [11], evaluates each processing module in isolation by providing it with reference inputs rather than the outputs of the preceding block in the pipeline. This approach ensures that the accuracy of each block is assessed independently, preventing error propagation from previous steps and allowing for a direct comparison between the estimated outputs and those obtained from the validated INDIP pipeline.

The full-pipeline validation, following the approach implemented by Kirk in [9], assesses the performance of the entire pipeline by considering the final output rather than intermediate steps. In this case, the pipeline is executed in its entirety, where each block processes the outputs of the preceding one, allowing errors to propagate naturally. This approach helps evaluate how inaccuracies at different stages influence the estimation of the walking speed, final output of the entire pipeline.

For both validation methods, the INDIP pipeline served as the reference for extracting DMOs. This system processes inertial data from a lower-back LB-MIMU while also integrating information from foot-mounted MIMUs and pressure insoles

to enhance the accuracy of gait event detection and spatio-temporal parameters estimation [49]. The validation consists of comparing DMOs extracted with MobGap from smartphone-based inertial data against DMOs extracted using INDIP from the INDIP sensor kit. This approach enables a thorough assessment of both individual algorithmic components and overall system performance, ensuring that smartphone-based gait analysis can reliably approximate gold-standard measurements.

To assess the validity and performance of the Mobilise-D pipeline, a combination of classification metrics and agreement measures was employed, depending on the evaluated block. The classification metrics used to evaluate GSD and ICD included accuracy, recall, precision, F1-score and specificity. These are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5.4}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{5.5}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5.6}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5.7}$$

where TP represents true positive elements, TN true negative elements, FP false positive elements, and FN false negative elements.

In addition to these classification metrics, agreement between the DMOs extracted from MobGap and those obtained from the INDIP pipeline was evaluated using the Intra-Class Correlation Coefficient (ICC). ICC values were interpreted as follows:

- < 0.5: Poor agreement
- $0.5 - 0.75$: Moderate agreement
- $0.75 - 0.9$: Good agreement
- > 0.9: Excellent agreement

To further assess absolute agreement, the absolute error, bias, and Limits of Agreement (LoA) were quantified for each walking bout, providing insight into the systematic differences between the two systems.

Absolute Error : Measures the absolute difference between the values obtained from the two systems. It is defined as:

$$AE = |DMO_{\text{MobGap}} - DMO_{\text{INDIP}}| \tag{5.8}$$

where DMO_{MobGap} and DMO_{INDIP} represent the digital mobility outcomes estimated by MobGap and INDIP, respectively.

Bias: Represents the mean difference between the two systems, indicating systematic overestimation or underestimation. It is computed as:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}(DMO_{\text{MobGap},i} - DMO_{\text{INDIP},i}) \tag{5.9}$$

where N is the number of walking bouts analyzed.

Limits of Agreement (LoA): Quantifies the range within which most differences between the two methods lie, offering an assessment of their agreement. LoA is calculated as:

$$\text{LoA} = \text{Bias} \pm 1.96 \cdot SD_{\text{diff}} \tag{5.10}$$

where SD_{diff} is the standard deviation of the differences between MobGap and INDIP estimates.

Additionally, relative errors were computed to evaluate the percentage difference between the DMOs extracted from smartphone-based inertial data and those obtained using the INDIP reference system. This is defined as:

$$\text{Relative Error} = \frac{|DMO_{\text{MobGap}} - DMO_{\text{INDIP}}|}{DMO_{\text{INDIP}}} \times 100 \tag{5.11}$$

which provides a normalized measure of deviation, allowing for a more meaningful comparison across different gait parameters.

These validation metrics offer a comprehensive evaluation of both the accuracy of gait event detection and the reliability of estimated mobility outcomes, ensuring that the smartphone-based analysis aligns with gold-standard reference measurements.

Gait sequence detection

The validation of GSD block was performed by classifying each sample as true positive, false positive, false negative, or true negative. From these classifications, accuracy, specificity, recall and precision were computed. Given that the inertial data was recorded at 100 Hz, this evaluation was conducted with a temporal resolution of 0.01 seconds, ensuring a fine-grained analysis of the detection process.

Following this sample-level validation, only gait sequences that overlapped with the reference GSD for at least 80% of their duration were considered for further

analysis. For these sequences, the duration in seconds was extracted and compared to the reference values. Bias and LoA, Mean Absolute Error (MAE), and ICC were computed to quantify agreement between the detected and reference gait sequences.

This validation approach ensures a comprehensive assessment of the GSD block, capturing both individual sample accuracy and overall sequence detection reliability against the gold-standard reference.

Initial contact detection

The validation of the ICD block was performed by identifying ICs within the gait sequences detected by the reference system. Each detected IC was classified as true positive, false positive, or false negative, allowing for the computation of recall, precision, and F1-score.

To establish the matching between detected ICs and reference ICs, a tolerance window of ± 0.5 seconds was centered around each reference IC. If a detected IC fell within this window, it was classified as a true positive. Conversely, detected ICs outside this window were considered false positives, while reference ICs without any corresponding detected IC within the window were classified as false negatives.

This matching process ensured that only temporally aligned ICs were validated against the ground truth, allowing for a robust assessment of the system's performance.

Additionally, the MAE in seconds was calculated to quantify the temporal discrepancy between detected (TP IC events) and reference IC events. This approach ensures a detailed evaluation of the ICD block, assessing both the accuracy of IC identification and the timing precision relative to the gold-standard reference.

Cadence estimation

The validation of the CAD block was conducted by comparing the estimated cadence values with those obtained from the reference system. The evaluation was performed using MAE to measure the direct difference between estimated and reference cadence, relative error to quantify the percentage deviation, and ICC to assess the level of agreement between the two systems.

Stride length estimation

The same performance metrics of Cadence were used to evaluate the SL block. MAE, relative error and ICC were computed also for this block.

Full Pipeline validation

To evaluate the performance of the entire pipeline, a set of agreement and error metrics was selected to assess how closely the walking speed aligned with one obtained from the reference system. The evaluation included ICC, Bias and LoA to measure systematic differences and the range within which most errors lie, absolute error to quantify the direct deviation, and relative error to express the discrepancy as a percentage. Additionally, relative absolute error was computed to provide a normalized measure of deviation.

These metrics ensure a comprehensive evaluation of the full pipeline by capturing both the accuracy of the final estimates and the extent to which errors propagate through the different processing blocks.

5.3 Smartphone location recognition

The process of smartphone location recognition follows the structured pipeline illustrated in Figure 5.2, consisting of two main phases: data preparation and model development and validation. These steps ensured that raw inertial data were properly processed and transformed into meaningful inputs for machine learning models.

The data preparation phase was conducted using *MATLAB* (version 2024b). The subsequent phase, involving model development and validation, was implemented in *Python* (version 3.12.0), utilizing its extensive machine learning libraries to train and evaluate classification models.

The following subsections will provide a detailed explanation of each step in the pipeline, covering the data processing techniques and classification strategies used in this study.

5.3.1 Dataset construction

The start and end points of each gait sequence identified by the INDIP pipeline were extracted to isolate valid gait sequences from the smartphone inertial data. Leveraging the temporal alignment between the LB-smartphone and the LB-MIMU, followed by the synchronization of all other smartphones with the LB-smartphone, it was possible to use the gait sequences detected by INDIP to segment and extract the corresponding accelerometer and gyroscope signals exclusively during walking periods for each smartphone. This ensures that only sequences truly representative of walking are included in the dataset.

To enhance dataset variability and improve the generalizability of the classification model, both laboratory trials and free-living recordings were incorporated.

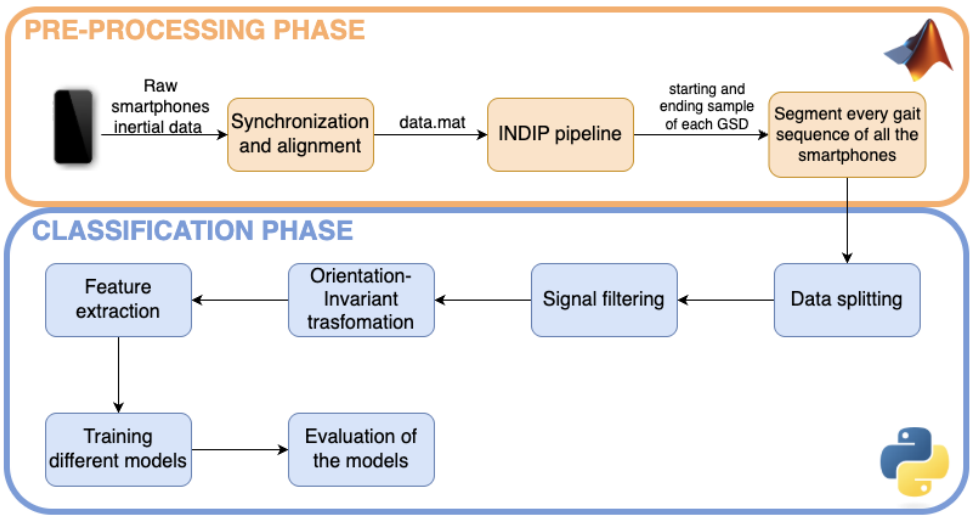


Figure 5.2: Overview of the smartphone location recognition pipeline. The process is divided into two main phases: Pre-Processing and Classification.

This allows the model to learn from diverse conditions, making it more robust across different walking scenarios.

As stated in Section 4.4.1, during free-living recordgins participants were allowed to remove the smartphones at their discretion, with the exception of the LB-smartphone, which remained fixed. As a result, while the gait sequences extracted from INDIP were guaranteed to correspond to actual walking periods for the LB-smartphone, this was not necessarily the case for the other smartphones. To prevent the inclusion of non-walking segments—such as instances where a smartphone was left on a desk or in a bag without movement—a quality control step was introduced.

A minimum gyroscope norm threshold was set to ensure that only walking sequences were included for analysis. Specifically, for a walking bout to be considered valid, the norm of the gyroscope signal had to exceed 1°/s for at least 80% of the detected sequence samples. This threshold, determined based on a empiric evaluation of various gait sequences, ensured that all extracted sequences corresponded exclusively to walking data.

An example of this filtering process is shown in Figure 5.3, where, for subject 17, the 8th walking bout detected by the INDIP pipeline was discarded for the H, SB, and CP smartphone positions.

As a result of this filtering process, a total of 412 walking bouts were discarded across all smartphone positions, with the highest number of discarded bouts observed for the ShoulderBag (146 bouts) and Hand (126 bouts) placements, as shown in Table 5.2. This confirms that in some positions, the smartphone was

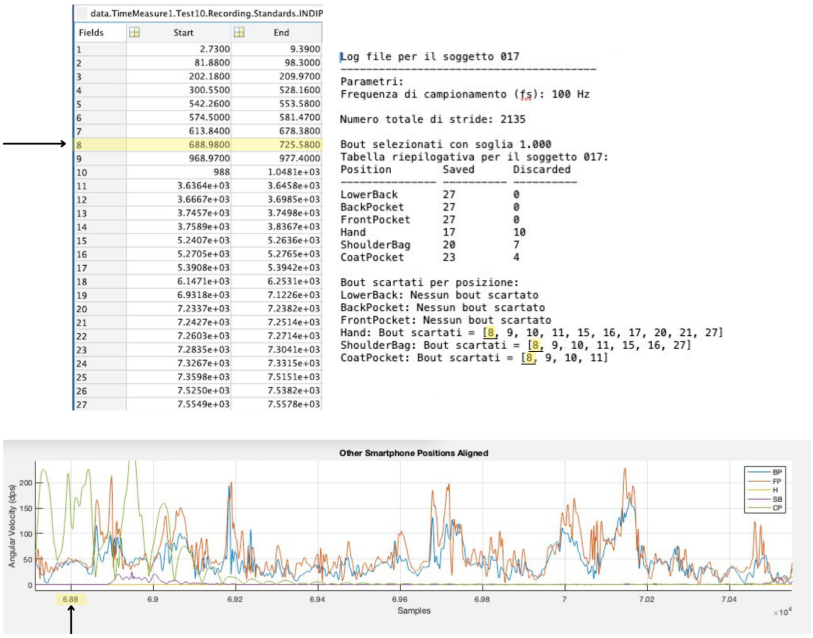


Figure 5.3: Example of a discarded gait sequence for subject 17. The 8th WB identified by the INDIP pipeline was removed for the Hand, ShoulderBag, and CoatPocket positions due to insufficient gyroscope activity. The top section of the figure illustrates the recorded walking bouts and their classification, while the lower plot shows the corresponding gyroscope signals, highlighting the insufficient motion detected for the discarded positions.

frequently removed or remained motionless during detected walking sequences, necessitating careful data selection to ensure only valid walking segments were included in the dataset.

Position	Saved	Discarded
LowerBack	645	0
BackPocket	621	24
FrontPocket	621	24
Hand	519	126
ShoulderBag	499	146
CoatPocket	522	92

Table 5.2: Number of walking bouts retained (Saved) and removed (Discarded) for each smartphone position after data selection.

This choice was made as part of the project design, where the classification task was structured as a supervised learning problem focused exclusively on identifying predefined smartphone positions. Alternatively, a different approach could have been taken by including additional classes, such as "non-walking" or "unknown position" leading to a broader classification task. However, such an approach would have introduced different challenges and objectives, diverging from the primary focus of this study.

5.3.2 Dataset preparation

Signal filtering

To isolate the gait-related components of the inertial signals while reducing high-frequency noise, a Butterworth low-pass filter was applied. This filter was chosen due to its maximally flat frequency response, which ensures a smooth attenuation of unwanted high-frequency components while preserving the primary motion characteristics of walking without excessive signal distortion. The magnitude and phase response of the filter, illustrated in Figure 5.4, confirm its effectiveness in suppressing unwanted frequency components while maintaining signal integrity.

A 4th-order Butterworth filter was implemented with a cutoff frequency normalized to the Nyquist frequency to effectively suppress noise while retaining gait-relevant signal information. However, standard filtering methods can introduce phase distortion, meaning that important gait-related events (such as initial contacts) could be shifted in time, potentially affecting the accuracy of subsequent feature extraction.

To avoid this issue, the zero-phase filtering method `filtfilt` was used. The Butterworth filter, being an infinite impulse response (IIR) filter, inherently introduces nonlinear phase shifts when applied in a single direction. `filtfilt` compensates for this by applying the filter twice—first forward and then backward—effectively canceling out the phase distortion introduced in each pass. This results in a zero-phase response, meaning that no time shifts are introduced in the filtered signal, while still benefiting from the smooth frequency response of the Butterworth filter.

The stability of the applied filter is confirmed by analyzing its pole-zero plot, shown in Figure 5.5. All poles of the designed filter lie inside the unit circle in the z-plane, demonstrating that the filter is stable and suitable for processing gait signals without introducing artifacts or numerical instabilities.

By applying zero-phase filtering, the temporal integrity of gait events is preserved, ensuring that stride timings, initial contacts, and other gait-related features remain accurate. This approach improves the reliability of subsequent feature extraction and classification processes, maintaining the physiological accuracy of the recorded movement data.

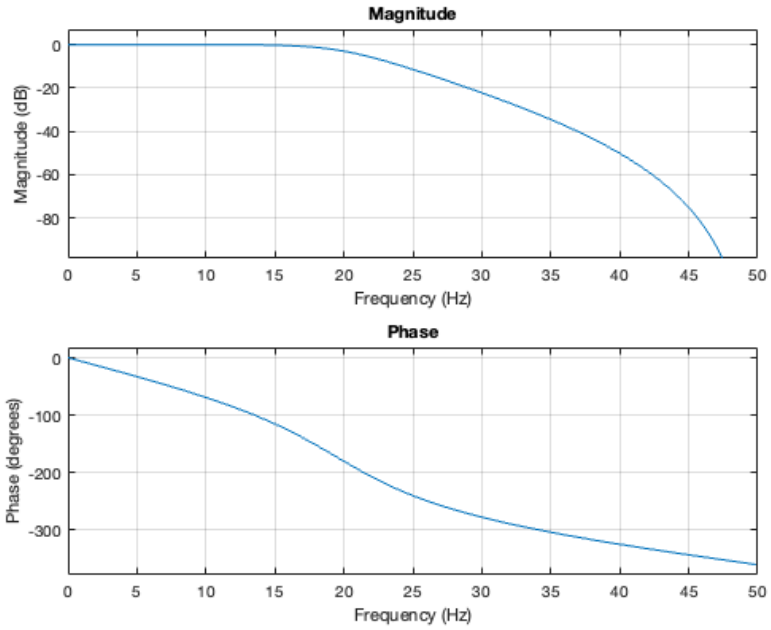


Figure 5.4: Magnitude and phase response of the implemented Butterworth low-pass filter. The filter effectively suppresses high-frequency noise while maintaining smooth signal attenuation.

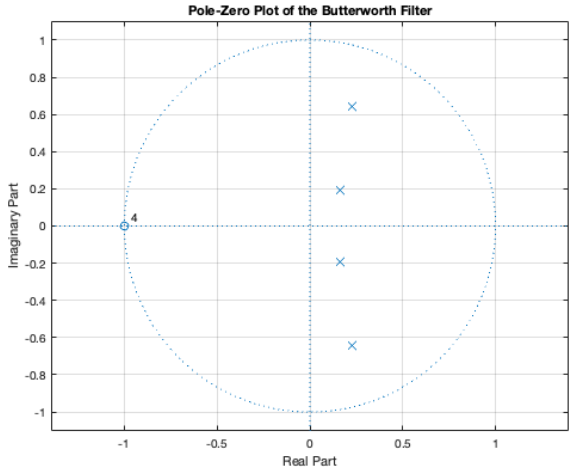


Figure 5.5: Pole-zero plot of the Butterworth filter. All poles lie inside the unit circle, confirming the stability of the filter.

Features extraction

To characterize gait patterns across different smartphone placements, a set of handcrafted features was extracted from accelerometer and gyroscope signals.

Except for the cross-correlation features, all features were computed on the signal norm to ensure orientation invariance (Orientation-Invariant approach, as illustrated in the pipeline Figure 5.2) minimizing the impact of sensor placement angles. The norm of a three-axis signal at sample i is defined as:

$$\|\mathbf{a}_i\| = \sqrt{x_i^2 + y_i^2 + z_i^2} \tag{5.12}$$

where x_i, y_i, z_i are the signal components along the three axes. This norm was used in all subsequent feature calculations.

Time-Domain Features Time-domain features describe the variability and distribution of the raw signals, providing insights into movement intensity, stability, and periodicity across different smartphone positions:

- **Variance**: Captures the spread of the signal values, representing movement intensity. It is defined as:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(\|\mathbf{a}_i\| - \bar{a})^2 \tag{5.13}$$

  where $\|\mathbf{a}_i\|$ represents the signal norm at sample i, $\bar{a}$ is the mean of the signal norm, and N is the total number of samples. Higher variance is associated with more dynamic motion (Observed for the Handheld smartphone), while lower variance suggests a more stable placement (e.g., Lower Back).

- **Root Mean Square (RMS)**: Provides a measure of the overall energy of the signal, incorporating both amplitude and frequency components. It is calculated as:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\|\mathbf{a}_i\|^2} \tag{5.14}$$

  Higher RMS values indicate greater motion intensity, differentiating highly dynamic and more stable smartphone placements.

- **Signal Magnitude Area (SMA)**: Computes the sum of absolute values of the signal norm, offering a robust measure of total movement intensity. It is defined as:

$$SMA = \frac{1}{N}\sum_{i=1}^{N}|\|\mathbf{a}_i\||| \tag{5.15}$$

  This feature provides a reliable measure of total movement intensity, particularly useful for identifying positions where the device experiences significant motion.

- **Number of Peaks**: Counts the local maxima in the signal, which reflects periodic movements. A higher number of peaks typically corresponds to increased gait variability or noise, often linked to less stable placements.
- **Cross-Correlation Between Axes**: Measures the relationship between different accelerometer and gyroscope axes (XY, XZ, YZ). It is computed as the Pearson correlation coefficient:

$$\rho_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{5.16}$$

  where x_i and y_i represent signal values from two different axes, and $\bar{x}, \bar{y}$ are their respective means. This feature provides information on movement synchronization and device orientation, helping distinguish constrained placements (e.g., Lower Back) from more freely moving positions (e.g., Hand).

Frequency-Domain Features Frequency-domain features were computed to capture periodic characteristics of gait signals, which are essential for identifying walking dynamics and differentiating smartphone placements:

- **Dominant Frequency**: Identifies the most prominent frequency component in the signal. It is obtained by applying the Fast Fourier Transform (FFT) and selecting the frequency with the highest power:

$$f_{\text{dom}} = \arg\max_f |X(f)| \tag{5.17}$$

  where $X(f)$ represents the magnitude of the signal in the frequency domain.
- **Dominant Power**: Measures the magnitude of the dominant frequency, indicating how periodic the movement is. It is defined as:

$$P_{\text{dom}} = \max |X(f)| \tag{5.18}$$

  where P_{dom} represents the peak power of the dominant frequency component. A strong dominant power suggests regular, structured motion (e.g., Lower Back), while a weaker dominant power indicates more irregular patterns (e.g., Hand).
- **Spectral Entropy**: Quantifies the complexity of the frequency content. It is computed as the Shannon entropy of the normalized power spectral density:

$$H = -\sum_i P_i \log P_i \tag{5.19}$$

  where P_i represents the power spectral density at frequency bin i. Higher entropy values suggest a broader frequency distribution, often associated with irregular movement, while lower entropy values indicate more periodic and structured motion.

Feature Extraction Strategy Features were extracted from each segmented walking bout using a sliding window approach to capture variations over different time scales:

- 1-second windows to capture short-term gait fluctuations.
- 3-second windows to encompass a broader view of gait cycles, assuming that in a healthy adult, at least one full gait cycle is completed within this duration, given a detected cadence of 85–90 steps per minute [11].

By combining time-domain and frequency-domain features, the classification model is provided with a rich representation of gait characteristics, improving its ability to differentiate smartphone positions based on their specific motion patterns.

5.3.3 Training and Validation

Data splitting

The dataset used for training and evaluating the smartphone location recognition model was constructed from 25 subjects recorded in Sheffield. However, due to issues such as insufficient signal quality or difficulties in gait sequence detection, only 15 subjects were retained for the final analysis. In total, 3.427 gait sequences (45551 strides) were included in the dataset.

As previously mentioned in the pipeline overview, different data splitting strategies were tested to assess model performance under various conditions. The first approach followed a train-test split using an 80/20 hold-out strategy. To prevent data leakage, the test set was not randomly selected at the walking bout level but rather at the subject level. Specifically, three subjects were isolated for testing, ensuring that the remaining 12 subjects in the training set provided a balanced distribution of walking bouts across all six smartphone positions. This approach was adopted to maintain a representative training set while preventing information from the same subject from appearing in both training and test sets.

In addition to the hold-out split, a 5-fold cross-validation was performed to further assess model robustness. In this approach, the dataset was divided into five subsets, and the model was trained and evaluated five times, each time using a different fold as the test set while training on the remaining four. This method provides a more comprehensive evaluation of the model's performance by reducing the dependence on a single train-test split.

Model selection

In the field of machine learning, especially for complex tasks such as smartphone location recognition, it is not possible to determine a priori which model will best

capture the necessary information to achieve optimal performance. Therefore, a result-driven approach was adopted, where multiple models were trained and evaluated empirically. A diverse set of machine learning models was selected to explore different learning paradigms and ensure a robust comparison.

The models considered for this study include:

- **Decision Tree**: A non-parametric model that recursively partitions the feature space by selecting the best splits based on Gini impurity, aiming to minimize classification uncertainty at each step.
- **Random Forest**: An ensemble of decision trees trained on randomly selected feature subsets and bootstrap samples, aggregating predictions via majority voting to improve generalization and reduce overfitting.
- **XGBoost**: An optimized gradient boosting algorithm where decision trees are trained sequentially, with each tree correcting the errors of the previous one. The boosting approach works by assigning higher weights to misclassified samples, making the model progressively focus on difficult cases while minimizing a differentiable loss function.
- **Logistic Regression**: A linear model that estimates class probabilities using the sigmoid function applied to a weighted sum of input features, making it effective for linearly separable classification problems.

Additionally, an **Artificial Neural Network (ANN)** was explored. The ANN architecture consisted of:

- An input layer with dimensionality corresponding to the number of extracted features.
- Three fully connected hidden layers with *128, 64, and 32 neurons*, respectively, each followed by a *ReLU activation function*.
- *Batch normalization* after the first two hidden layers to stabilize training and improve convergence.
- *Dropout regularization* with a dropout rate of 0.3 after the first two hidden layers to prevent overfitting.
- A final output layer with a *softmax* activation function for multi-class classification and a *sigmoid* activation function for the binarized problem.

The ANN model was optimized using the *Adam optimizer* with a learning rate of 0.005 and trained with a *sparse categorical cross-entropy* loss function for the multiclass problem. For the binarized problem a *binary cross-entropy* was

implemented as loss function. Training was performed for a maximum of 50 epochs, with early stopping enabled and a patience of 10 epochs to prevent overfitting. The batch size was set to 128, as it provides a good balance between computational efficiency and model generalization. A larger batch size accelerates training by allowing more parallel computations, while a smaller batch size introduces more stochasticity, which can help escape local minima.

By training and evaluating these models, the goal was to determine the most effective approach for smartphone location recognition, balancing model complexity, generalization ability, and computational efficiency.

5.3.4 Testing

The classification problem was approached in two different ways to evaluate the robustness of the models in distinguishing smartphone positions. The first approach considered a *multi-class classification*, where the model attempted to differentiate between all six predefined positions. The second approach used a *binary classification* strategy, where the goal was to distinguish between the LB placement and all other positions combined.

The decision to include the binary classification focusing on the LB placement was motivated by its relevance in gait analysis. Although the LB position is not commonly used in daily-life smartphone carrying habits, it is a well-established reference point for gait assessment. The *Mobilise-D pipeline* was designed and validated specifically using LB-mounted wearable devices, making it a key benchmark position in digital mobility outcome (DMO) extraction. Therefore, evaluating whether a model could reliably detect whether the smartphone was placed in this reference position was of particular interest.

To assess model performance, the following evaluation metrics were considered:

- **Confusion Matrices**: Provide a detailed breakdown of correct and incorrect classifications for each class, offering insights into common misclassifications. The confusion matrix for the multi-class classification problem is shown in Figure 5.6, while the confusion matrix for the binary classification approach (LB vs. others) is illustrated in Figure 6.40.

- **Accuracy**: Measures the overall proportion of correctly classified instances over the total number of instances and is defined as:

$$\text{Accuracy} = \frac{\sum_{i=1}^{C} TP_i}{\sum_{i=1}^{C} (TP_i + FP_i + FN_i)} \tag{5.20}$$

  where TP_i represents the true positives for class i, FP_i the false positives, FN_i the false negatives, while C is the total number of classes.

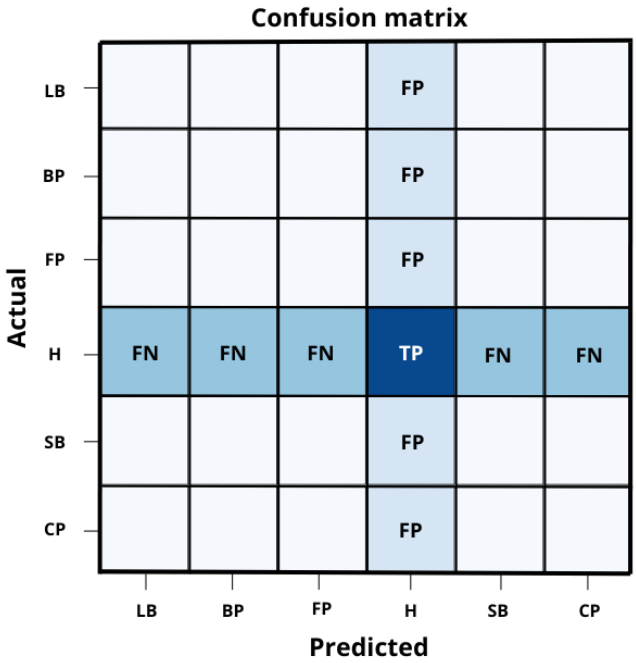


Figure 5.6: Example of a confusion matrix for the multi-class classification problem.

- **Balanced Accuracy**: Accounts for class imbalance by computing the average recall across all classes, ensuring a fair evaluation when class distributions are uneven. It is defined as:

$$\text{Balanced Accuracy} = \frac{1}{C} \sum_{i=1}^{C} \frac{TP_i}{TP_i + FN_i} \tag{5.21}$$

  where C is the number of classes, and the recall for each class is computed separately before averaging.

- **Precision**: Indicates how many instances classified as a specific position were actually correct, reflecting the reliability of positive predictions. It is defined as:

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} \tag{5.22}$$

- **Recall**: Measures the ability of the model to correctly identify instances of each class, assessing how well each smartphone placement is detected. It is

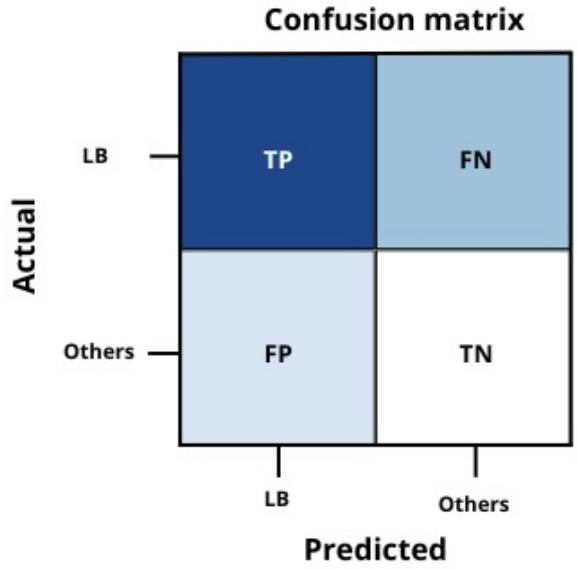


Figure 5.7: Example of a confusion matrix for the binary classification problem (LB vs. others).

computed as:

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i} \tag{5.23}$$

Chapter 6

Results

6.1 Mobilise-D Pipeline

This section presents the results of the validation of the Mobilise-D pipeline. The results are structured into two parts: first, the performance of each individual module within the pipeline is provided separately, providing a block-by-block evaluation. Then, to assess the overall effectiveness of the pipeline, the final output is considered, focusing on the estimation of walking speed. The results include both data collected in controlled laboratory conditions and in free-living scenarios. Performance metrics are reported for data acquired from the smartphone positioned at the lower back as well as from the MIMU of the INDIP kit.

6.1.1 Gait Sequence Detection

In-Lab: Performance Metrics

The following figures illustrate the distributions of the performance metrics obtained from smartphone-acquired data in a controlled laboratory setting. The results refer to the different algorithms available in MobGap for the GSD module. The box plots represent the variability of key metrics, including accuracy, precision, recall, specificity, total duration absolute error, and total duration bias, providing an overview of the performance across the available methods.

Free-living: Performance Metrics

The following figures instead, present the distributions of the performance metrics obtained from smartphone-acquired data in free-living conditions.

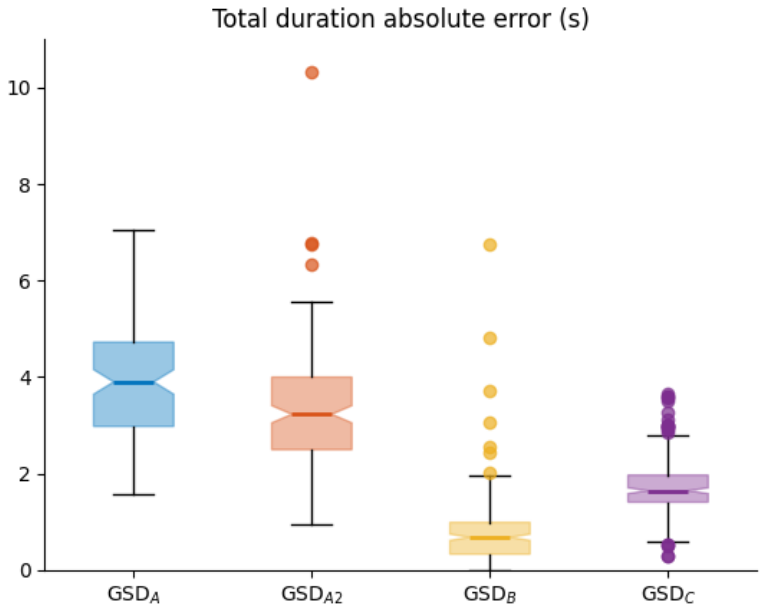


Figure 6.1: Total duration absolute error (s) for all algorithms in the GSD module using laboratory-acquired data in MobGap.

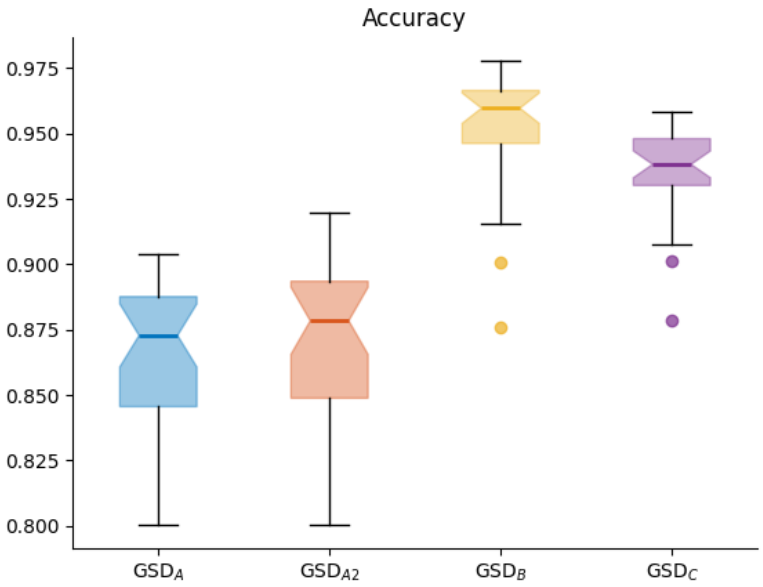


Figure 6.2: Accuracy distribution for all algorithms in the GSD module using laboratory-acquired data in MobGap.

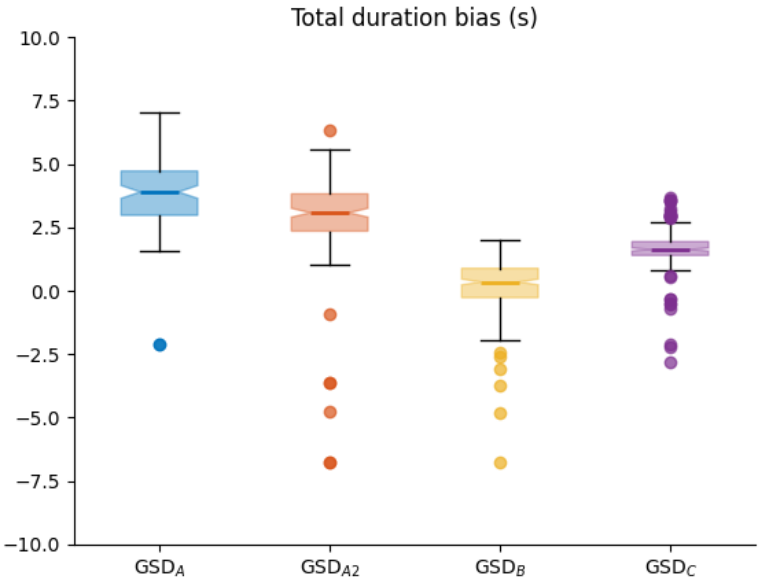


Figure 6.3: Total duration bias (s) for all algorithms in the Gait Sequence Detection (GSD) module using laboratory-acquired data in MobGap.

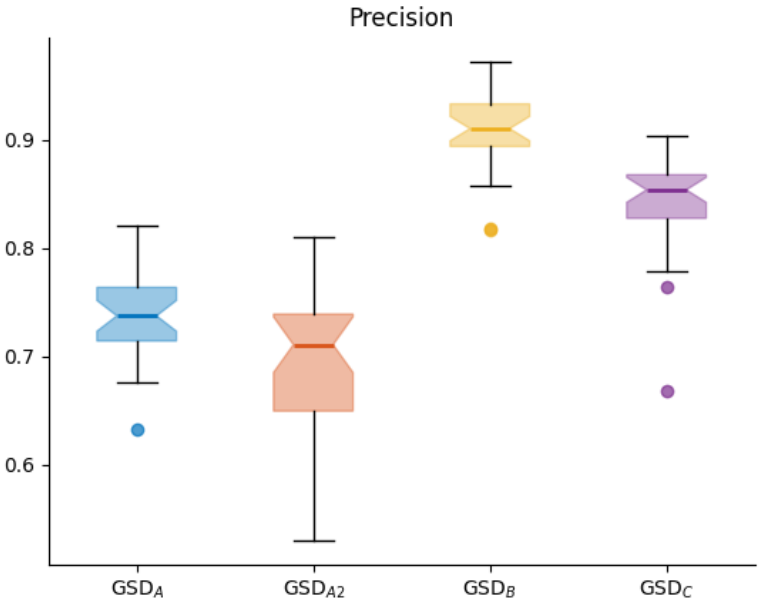


Figure 6.4: Precision distribution for all algorithms in the Gait Sequence Detection (GSD) module using laboratory-acquired data in MobGap.

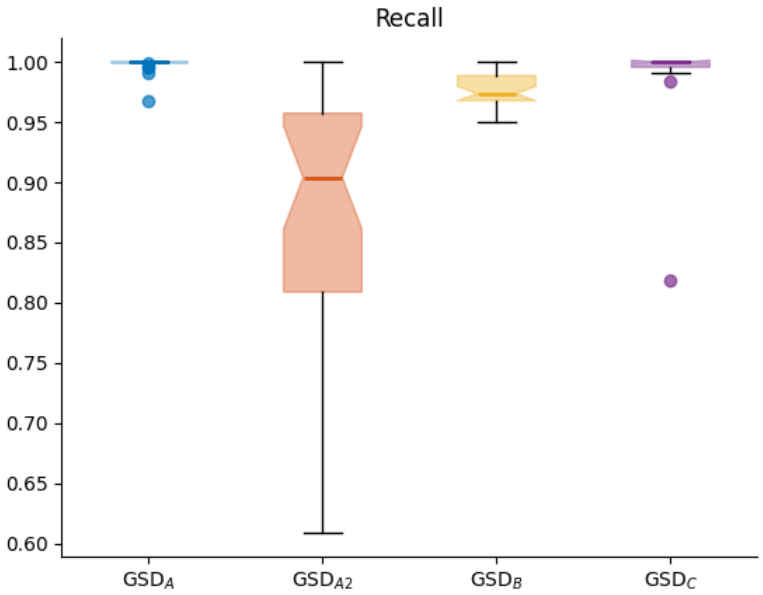


Figure 6.5: Recall distribution for all algorithms in the GSD module using laboratory-acquired data in MobGap.

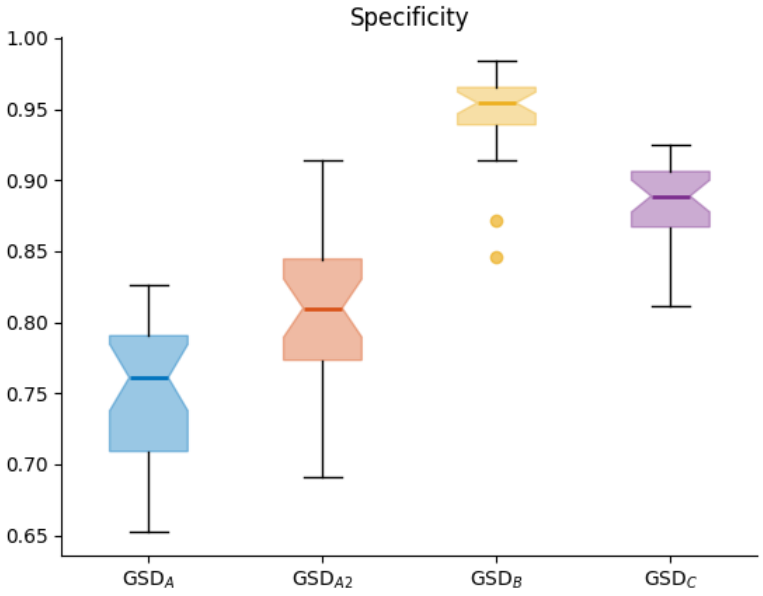


Figure 6.6: Specificity distribution for all algorithms in the GSD module using laboratory-acquired data in MobGap.

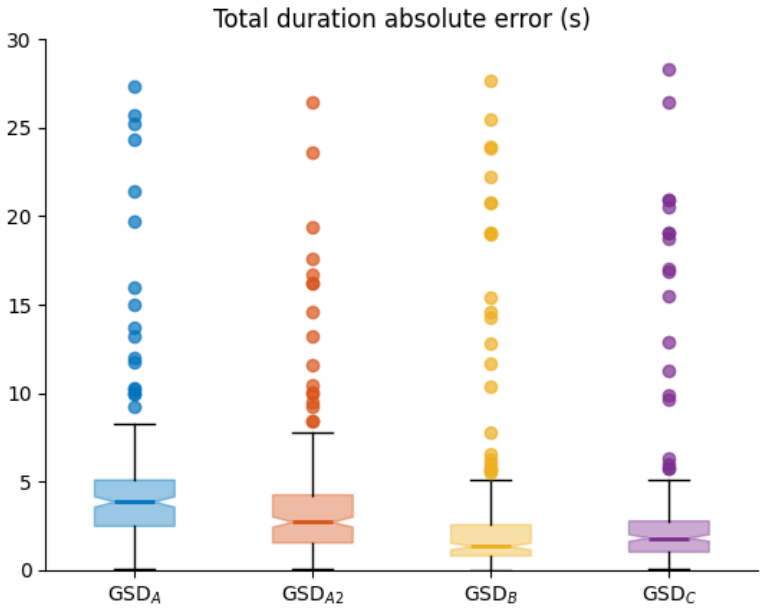


Figure 6.7: Total duration absolute error (s) for all algorithms in the GSD module using free-living data in MobGap.

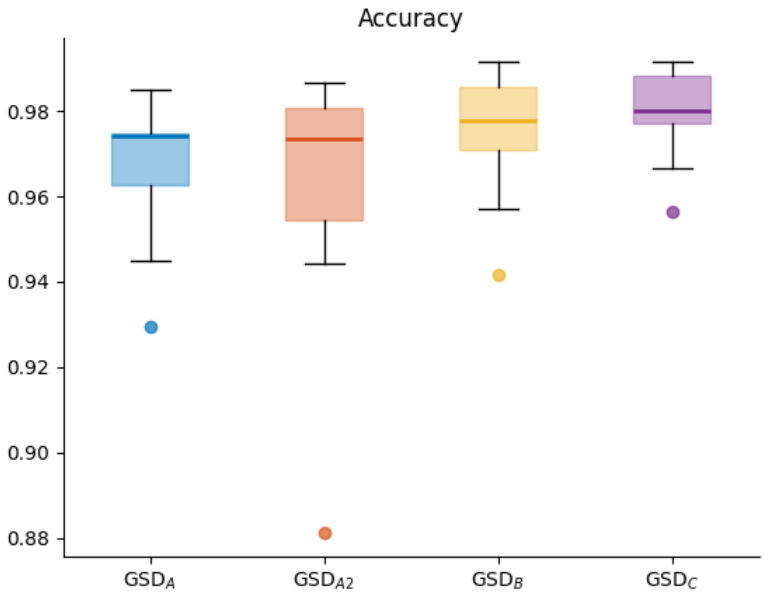


Figure 6.8: Accuracy distribution for all algorithms in the GSD module using free-living data in MobGap.

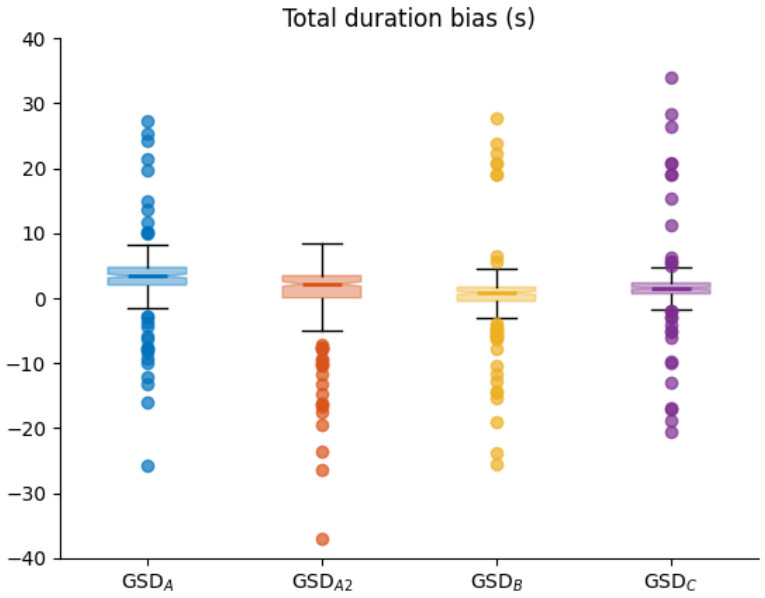


Figure 6.9: Total duration bias (s) for all algorithms in the GSD module using free-living data in MobGap.

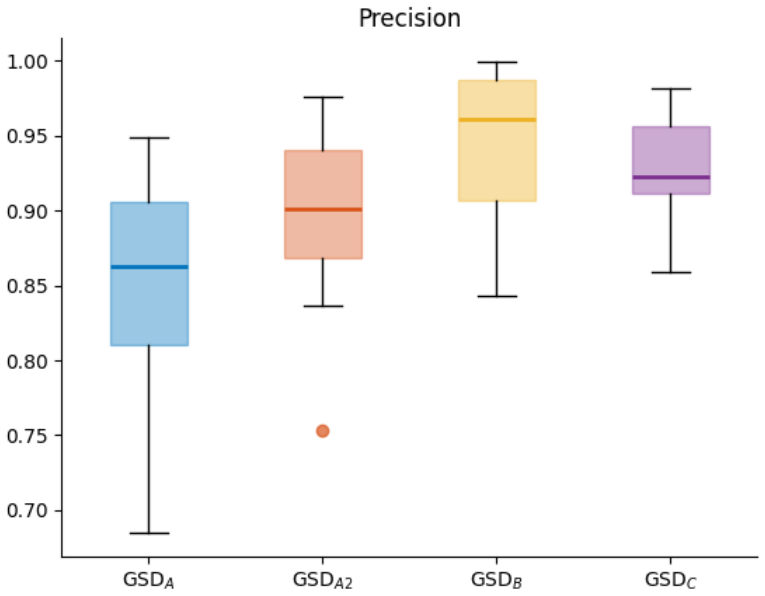


Figure 6.10: Precision distribution for all algorithms in the GSD module using free-living data in MobGap.

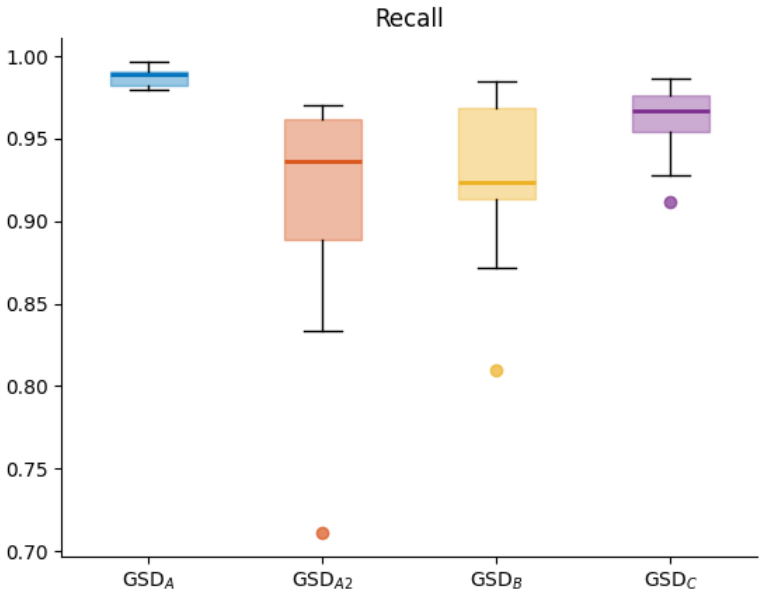


Figure 6.11: Recall distribution for all algorithms in the GSD module using free-living data in MobGap.

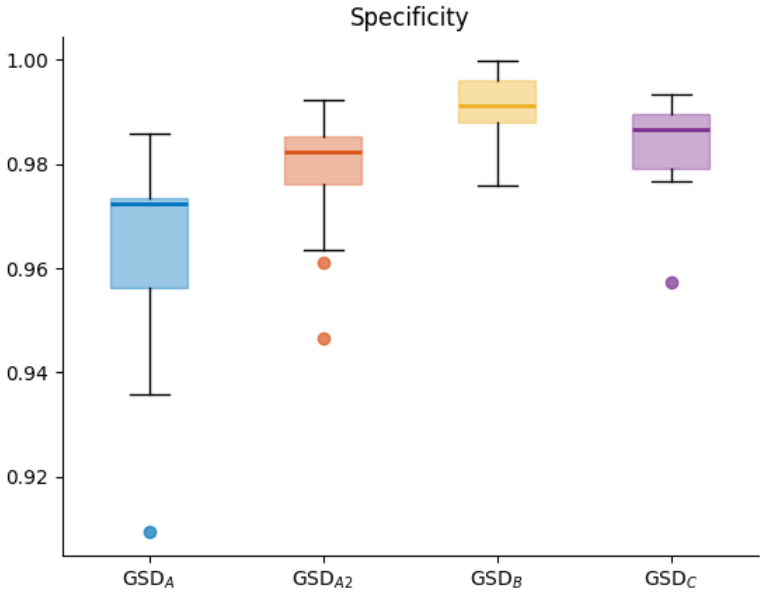


Figure 6.12: Specificity distribution for all algorithms in the GSD module using free-living data in MobGap.

Comparison: Smartphone vs IMU

The following table presents a comparative analysis of the GSD module, highlighting the differences between data collected using a smartphone and an INDIP IMU. The evaluation is conducted in both laboratory and free-living conditions. The table includes key performance indicators such as accuracy, precision, recall, specificity, total duration absolute error, and total duration bias, along with their confidence. This comparison provides insights into the impact of the sensing device on the performance of the detection algorithms.

	In Lab	
	Smartphone	**IMU**
Accuracy	0.87	0.87
Precision	0.74	0.74
Recall	0.99	0.99
Specificity	0.75	0.74
Duration absolute error (s)	3.86 [1.00, 10.50]	3.95 [1.10, 11.00]
Duration bias (s)	3.79 [-6.50, 12.00]	3.65 [-7.00, 11.50]
ICC	0.99	0.99

	Free-living	
	Smartphone	**IMU**
Accuracy	0.97	0.97
Precision	0.85	0.85
Recall	0.99	0.98
Specificity	0.96	0.96
Duration absolute error (s)	4.78 [1.05, 12.70]	5.40 [0.62, 21.99]
Duration bias (s)	3.22 [-7.84, 14.27]	3.17 [-12.68, 19.03]
ICC	0.99	0.99

Table 6.1: Comparison of gait sequence detection performance for smartphone and IMU in both lab and free-living conditions, including confidence intervals for error metrics.

6.1.2 Initial Contact Detection

In-Lab: Performance Metrics

The following figures illustrate the distributions of the performance metrics obtained from smartphone-acquired data in a controlled laboratory setting. The results refer to the different algorithms available in MobGap for the ICD module. The box plots represent the variability of key metrics, including absolute error, F1-score, precision, and recall, providing an overview of the performance across the available methods.

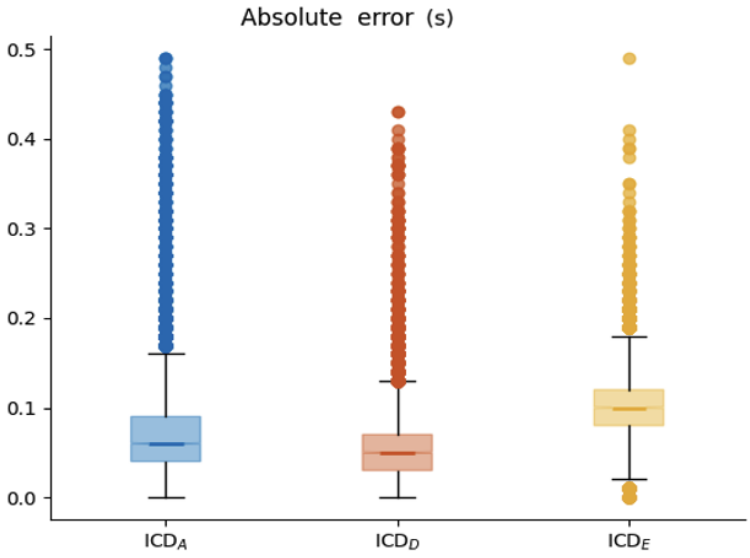


Figure 6.13: Absolute error (s) for all algorithms in the ICD module using laboratory-acquired data in MobGap.

Free-living: Performance Metrics

The following figures instead, present the distributions of the performance metrics obtained from smartphone-acquired data in free-living conditions.

Comparison: Smartphone vs IMU

The following table presents a comparative analysis of the ICD module, highlighting the differences between data collected using a smartphone and an INDIP IMU. The evaluation is conducted in both laboratory and free-living conditions.

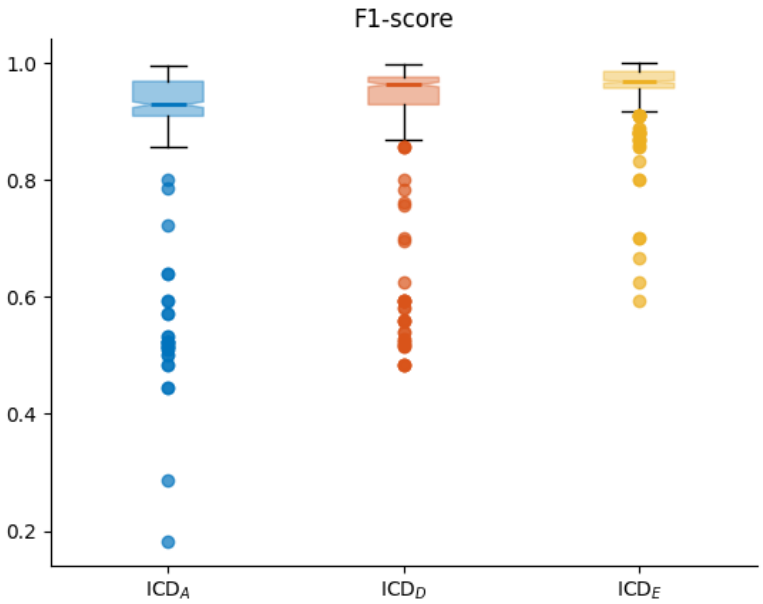


Figure 6.14: F1-score distribution for all algorithms in the ICD module using laboratory-acquired data in MobGap.

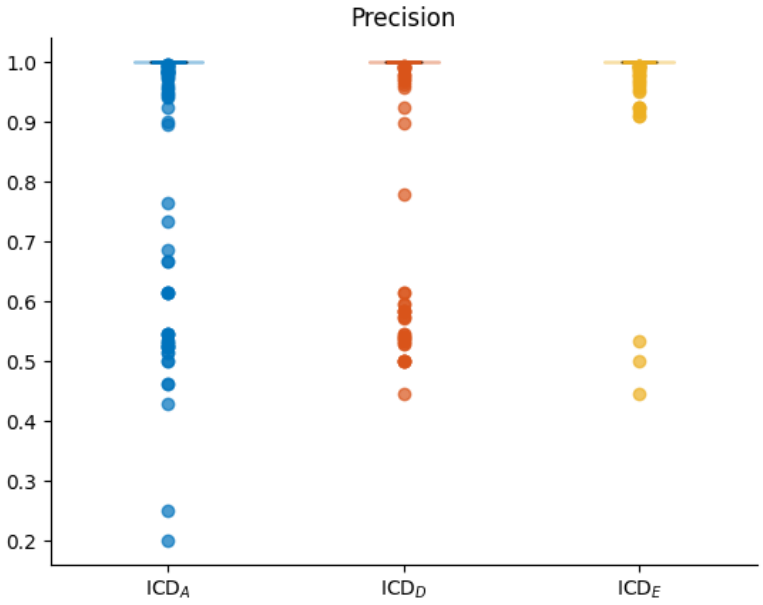


Figure 6.15: Precision distribution for all algorithms in the ICD module using laboratory-acquired data in MobGap.

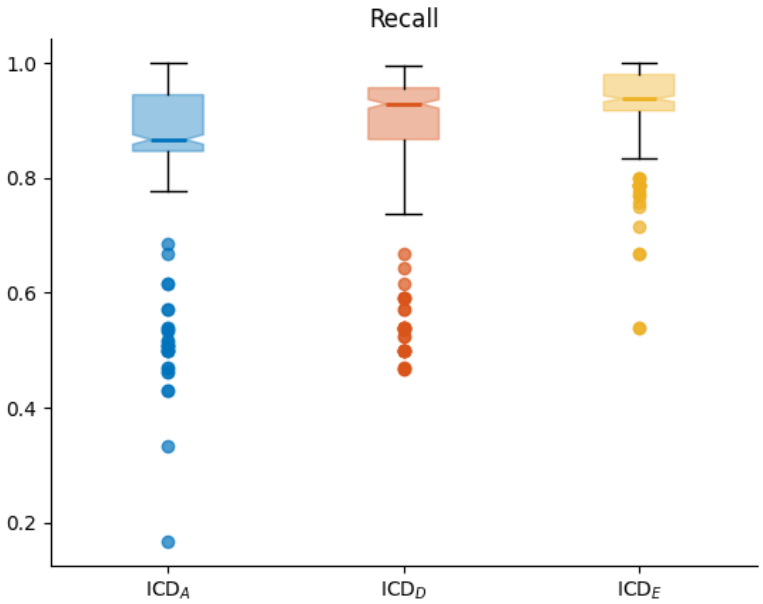


Figure 6.16: Recall distribution for all algorithms in the ICD module using laboratory-acquired data in MobGap.

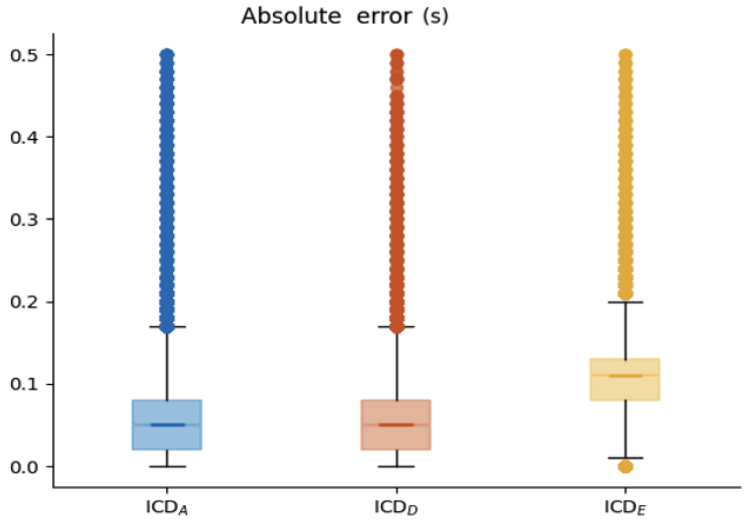


Figure 6.17: Absolute error (s) for all algorithms in the ICD module using free-living data in MobGap.

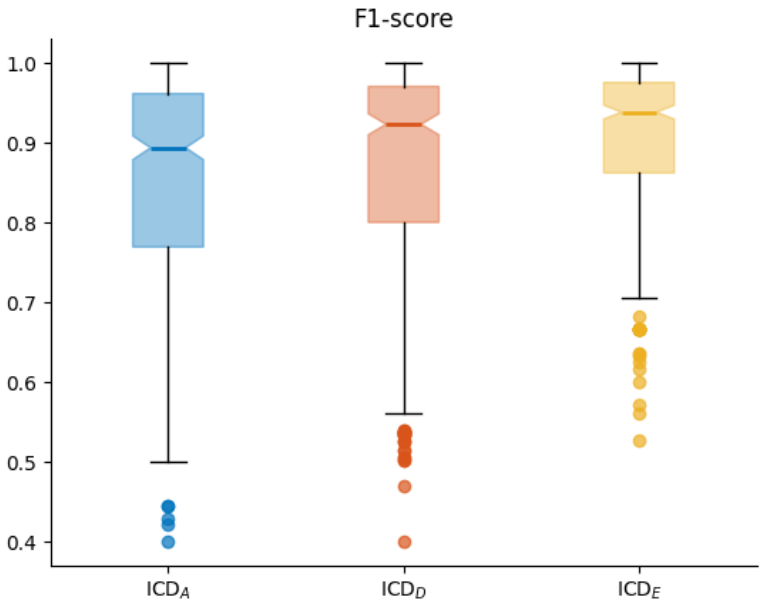


Figure 6.18: F1-score distribution for all algorithms in the ICD module using free-living data in MobGap.

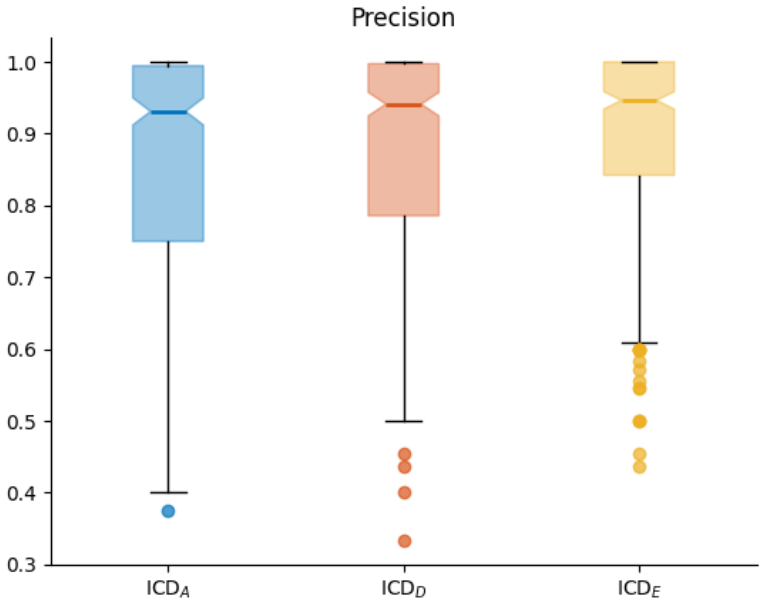


Figure 6.19: Precision distribution for all algorithms in the ICD module using free-living data in MobGap.

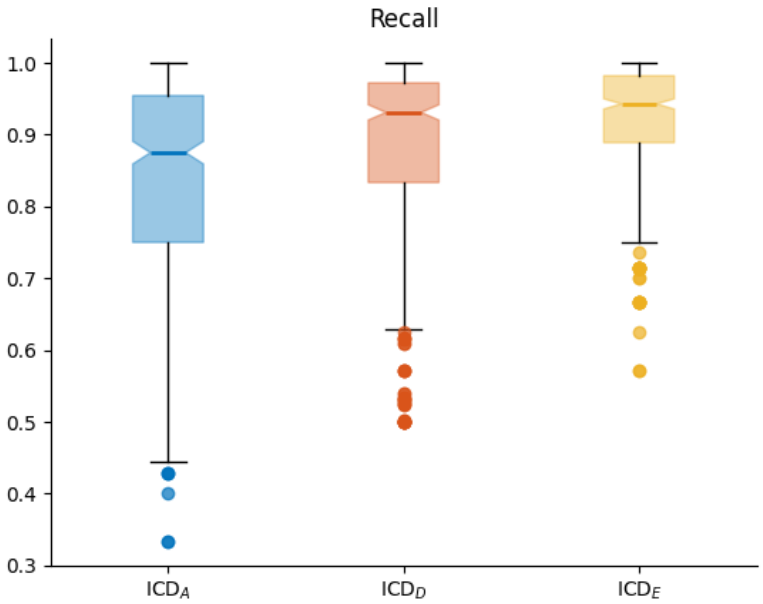


Figure 6.20: Recall distribution for all algorithms in the ICD module using free-living data in MobGap.

	In Lab	
	Smartphone	**INDIP**
F1-score	0.904	0.906
Precision	0.956	0.958
Recall	0.860	0.862
Absolute error (s)	0.069 [0.01, 0.18]	0.069 [0.01, 0.17]

	Free-living	
	Smartphone	**INDIP**
F1-score	0.842	0.838
Precision	0.862	0.859
Recall	0.832	0.828
Absolute error (s)	0.063 [0.01, 0.17]	0.065 [0.01, 0.17]

Table 6.2: Comparison of Initial Contact Detection (ICD) performance metrics between Smartphone and INDIP in both lab and free-living conditions, including confidence intervals for absolute error.

6.1.3 Cadence Estimation

In-Lab: Performance Metrics

The following figures illustrate the distributions of the performance metrics obtained from smartphone-acquired data in a controlled laboratory setting. The results refer to the different algorithms available in MobGap for the CAD module. The box plots represent the variability of key metrics, including absolute error and relative error providing an overview of the performance across the available methods.

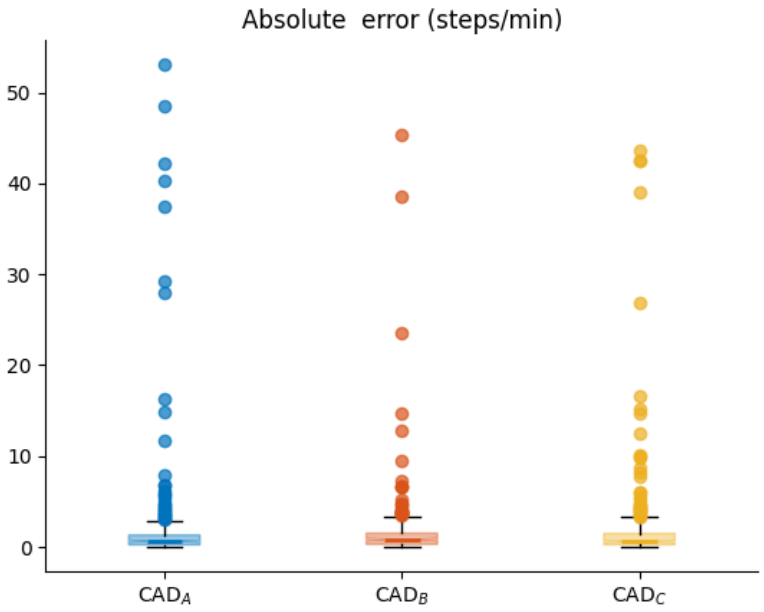


Figure 6.21: Absolute error (steps/min) for all algorithms in the CAD module using laboratory-acquired data in MobGap.

Free-living: Performance Metrics

The following figures instead, present the distributions of the performance metrics obtained from smartphone-acquired data in free-living conditions.

Comparison: Smartphone vs IMU

The following table presents a comparative analysis of the CAD module, highlighting the differences between data collected using a smartphone and an INDIP IMU. The evaluation is conducted in both laboratory and free-living conditions.

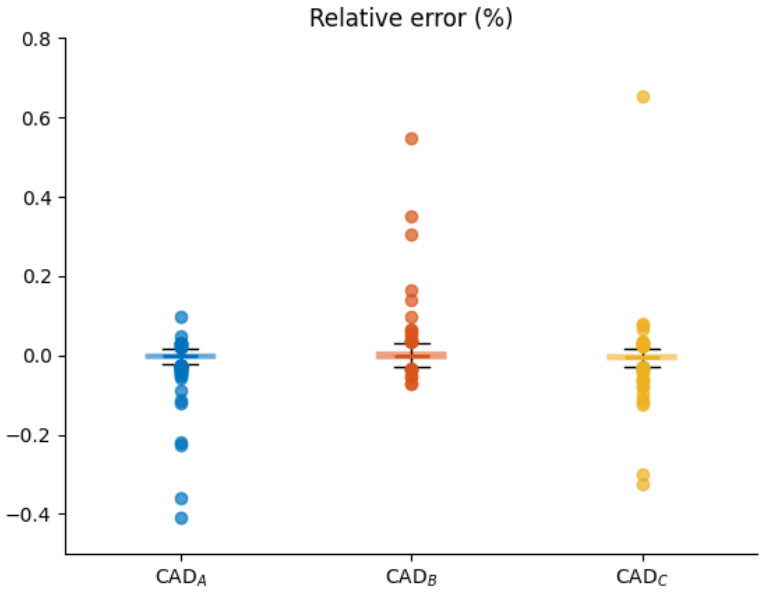


Figure 6.22: Relative error (%) for all algorithms in the CAD module using laboratory-acquired data in MobGap.

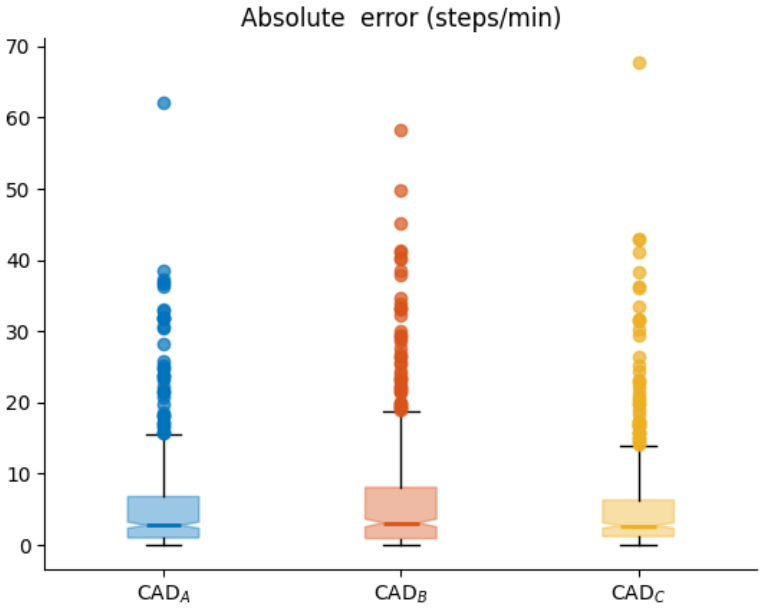


Figure 6.23: Absolute error (steps/min) for all algorithms in the CAD module using free-living data in MobGap.

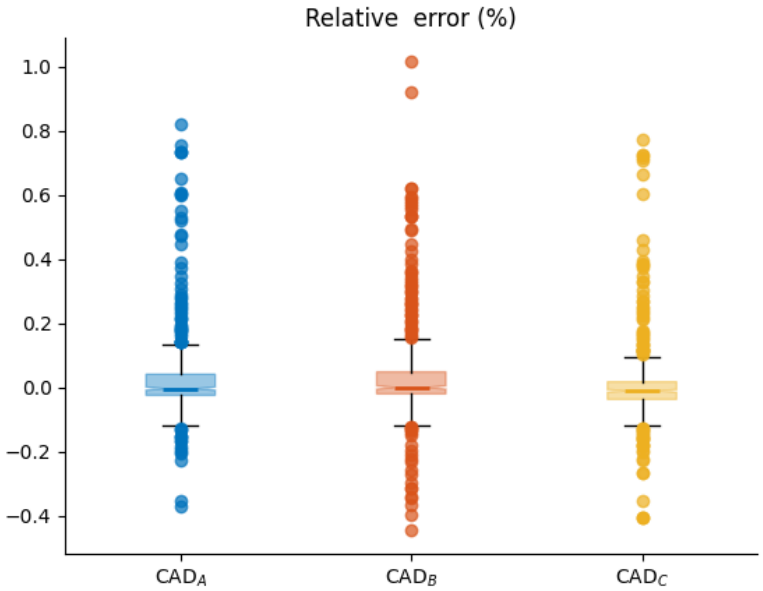


Figure 6.24: Relative error (%) for all algorithms in the CAD module using free-living data in MobGap.

	In Lab	
	Smartphone	**INDIP**
Absolute error (steps/min)	1.86 [0.07, 5.70]	1.75 [0.05, 5.23]
Relative error (%)	0.00 [-0.18, 0.18]	0.01 [-0.17, 0.18]

	Free-living	
	Smartphone	**INDIP**
Absolute error (steps/min)	5.66 [0.28, 21.08]	5.77 [0.28, 21.96]
Relative error (%)	0.01 [-0.26, 0.27]	0.01 [-0.26, 0.28]

Table 6.3: Comparison of Cadence Detection (CAD) performance metrics between Smartphone and INDIP in both lab and free-living conditions, including confidence intervals for error metrics.

6.1.4 Stride Length Estimation

In-Lab: Performance Metrics

The following figures illustrate the distributions of the performance metrics obtained from smartphone-acquired data in a controlled laboratory setting. The results refer

to the different algorithms available in MobGap for the SL module. The box plots represent the variability of key metrics, including absolute error, relative error and bias providing an overview of the performance across the available methods.

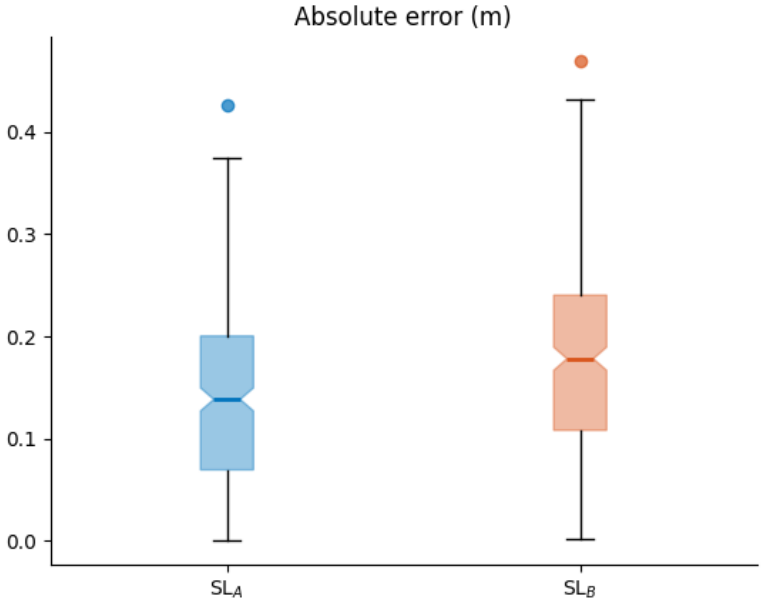


Figure 6.25: Absolute error (m) for all algorithms in the SL module using laboratory-acquired data in MobGap.

Free-living: Performance Metrics

The following figures instead, present the distributions of the performance metrics obtained from smartphone-acquired data in free-living conditions.

Comparison: Smartphone vs IMU

The following table presents a comparative analysis of the SL module, highlighting the differences between data collected using a smartphone and an INDIP IMU. The evaluation is conducted in both laboratory and free-living conditions.

6.1.5 Full Pipeline: Walking Speed

Comparison: Smartphone vs IMU

The following section presents a comparative analysis of the Walking Speed (WS) estimation obtained from smartphone-acquired data and IMU data in both laboratory and free-living conditions. The results, summarized in Table 6.5, include key

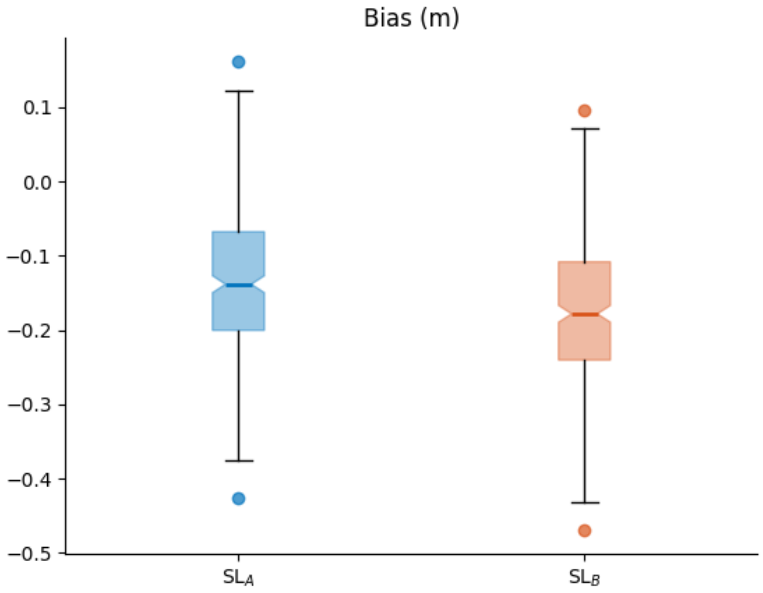


Figure 6.26: Bias (m) for all algorithms in the SL module using laboratory-acquired data in MobGap.

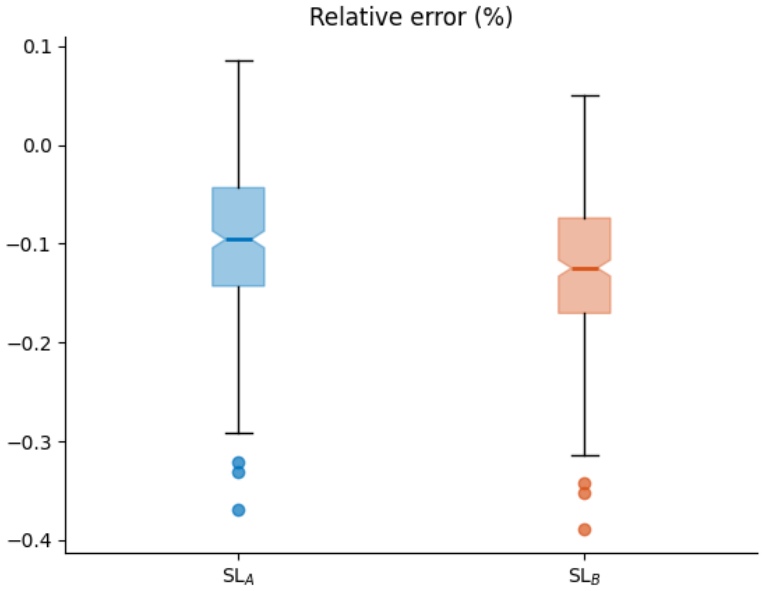


Figure 6.27: Relative error (%) for all algorithms in the SL module using laboratory-acquired data in MobGap.

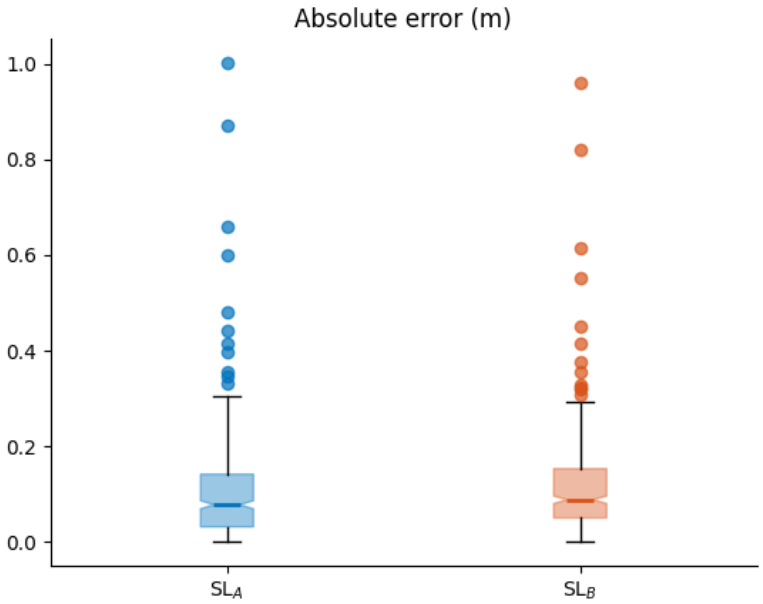


Figure 6.28: Absolute error (m) for all algorithms in the SL module using free-living data in MobGap.

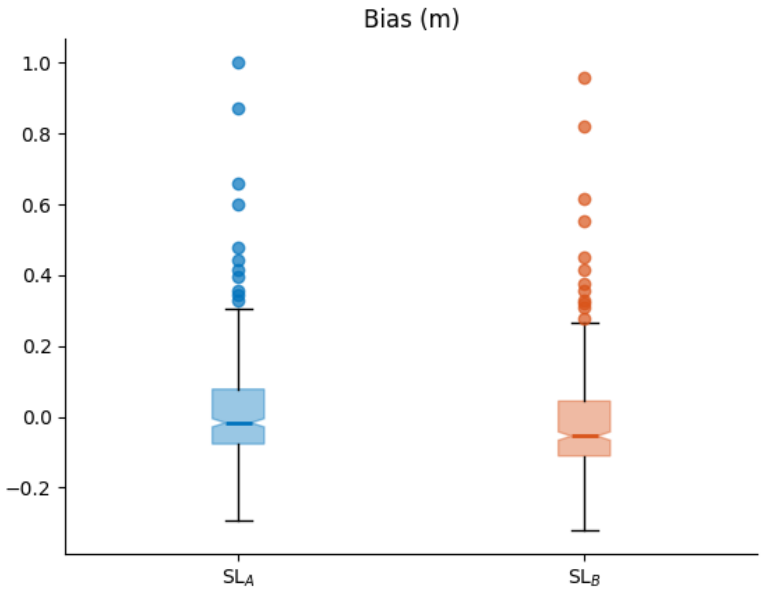


Figure 6.29: Bias (m) for all algorithms in the SL module using free-living data in MobGap.

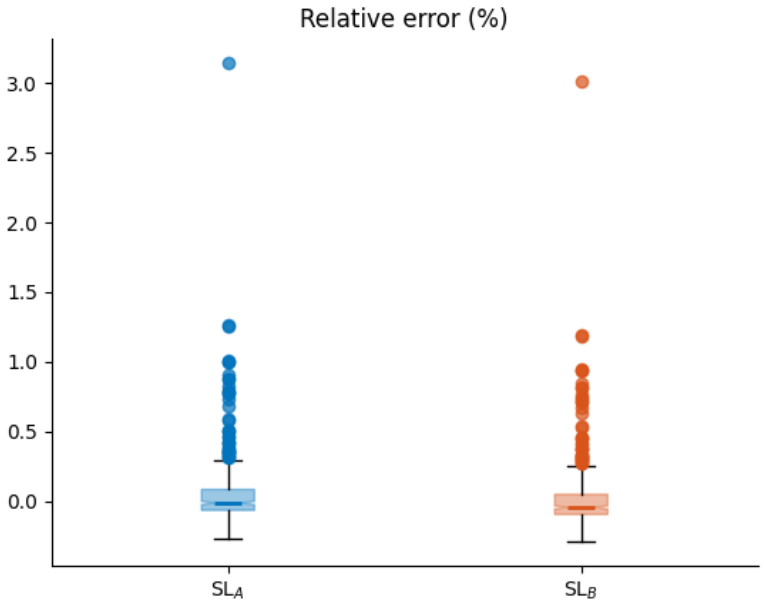


Figure 6.30: Relative error (%) for all algorithms in the SL module using free-living data in MobGap.

	In Lab	
	Smartphone	**INDIP**
Absolute error (m)	0.17 [0.03, 0.33]	0.18 [0.03, 0.35]
Relative error (%)	-0.13 [-0.27, 0.01]	-0.13 [-0.27, 0.02]
Bias (m)	-0.18 [-0.35, -0.03]	-0.18 [-0.35, -0.01]

	Free-living	
	Smartphone	**INDIP**
Absolute error (m)	0.11 [0.01, 0.26]	0.12 [0.01, 0.28]
Relative error (%)	0.02 [-0.48, 0.53]	0.03 [-0.49, 0.54]
Bias (m)	-0.02 [-0.20, 0.23]	-0.02 [-0.22, 0.24]

Table 6.4: Comparison of Step Length (SL) performance metrics between Smartphone and INDIP in both lab and free-living conditions, including confidence intervals for error metrics.

performance metrics such as absolute error, relative error, and bias, with confidence intervals. This comparison highlights the influence of different sensing devices on

the performance of the Mobilise-D pipeline in estimating walking speed.

In addition to the numerical comparison, Figures 6.31 and 6.32 illustrate Bland-Altman plots assessing the agreement between the walking speed values computed using the smartphone and those obtained from the IMU. For these plots, the median walking speed was selected from all computed values for the walking bouts identified for each subject. The Bland-Altman analysis allows for a visual inspection of systematic differences, potential biases, and agreement limits between the two sensor types.

	In Lab	
	Smartphone	**IMU**
Absolute error (m/s^2)	0.15 [0.04, 0.27]	0.16 [0.03, 0.31]
Relative error (%)	-0.15 [-0.29, 0.00]	-0.15 [-0.31, 0.00]
Bias (m/s^2)	-0.15 [-0.29, 0.00]	-0.16 [-0.31, 0.00]

	Free-living	
	Smartphone	**IMU**
Absolute error (m/s^2)	0.09 [0.01, 0.23]	0.09 [0.01, 0.21]
Relative error (%)	0.00 [-0.44, 0.43]	-0.03 [-0.35, 0.29]
Bias (m/s^2)	-0.04 [-0.25, 0.17]	-0.04 [-0.25, 0.16]

Table 6.5: Comparison of Walking Speed (WS) performance metrics between Smartphone and IMU in both lab and free-living conditions, including confidence intervals for error metrics.

6.2 Smartphone Location Recognition Model

This section presents the results of the smartphone location recognition model across different classification tasks. The primary analysis was conducted on a *6-classes problem*, where the model aimed to distinguish among all six predefined smartphone placements. As an extension of this analysis, a *5-classes problem* was introduced to explore an alternative class grouping strategy and assess its impact on classification performance. Additionally, a *binary classification problem* was explored to assess the model's ability to differentiate between the lower back (LB) placement and all other positions, given its relevance in gait analysis and its role in the *Mobilise-D pipeline*.

In the next sections, only the results obtained using the *5-Folds Cross-Validation*

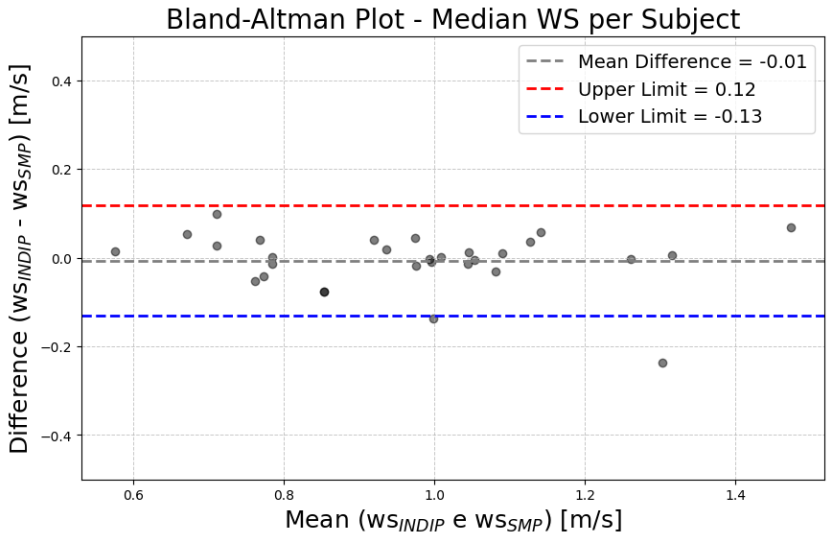


Figure 6.31: Bland-Altman plot comparing walking speed estimation between Smartphone and IMU in laboratory conditions. The median walking speed per subject was used for this analysis.

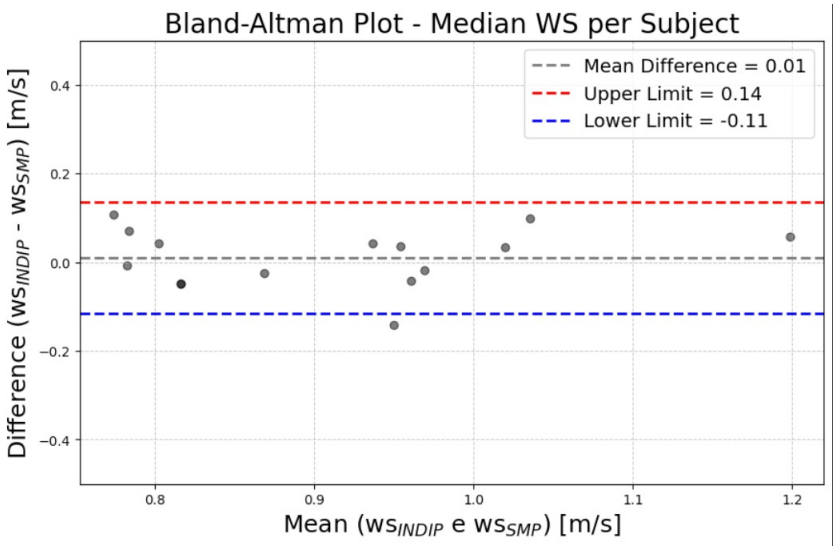


Figure 6.32: Bland-Altman plot comparing walking speed estimation between Smartphone and IMU in free-living conditions. The median walking speed per subject was used for this analysis.

(CV) strategy are presented. This choice is motivated by the fact that cross-validation provides a more reliable and robust assessment of model performance

compared to the *80-20 Hold-Out* approach. The latter, relying on a single data split, may introduce a positive bias, potentially overestimating the model's generalizability.

6.2.1 6-classes Classification Problem

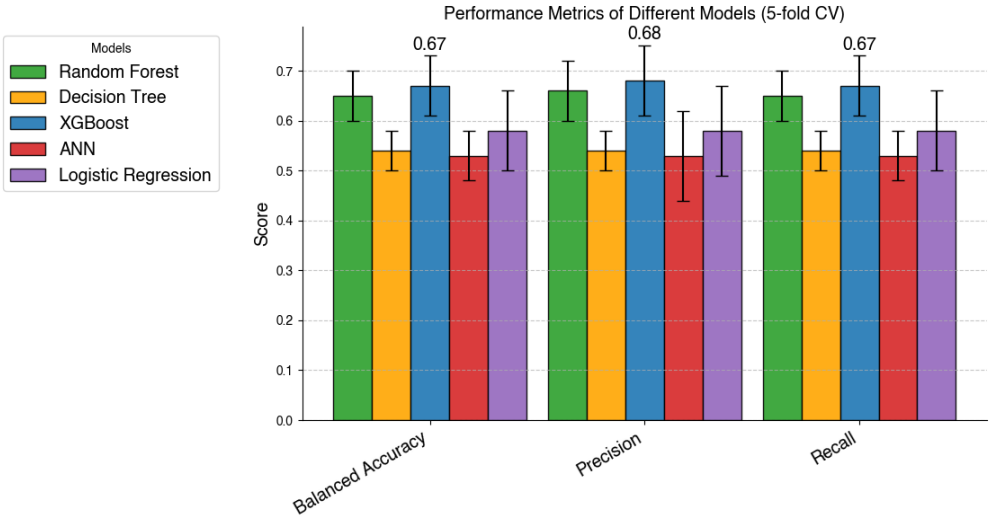


Figure 6.33: Performance metrics (Balanced Accuracy, Precision, Recall) for different classification models evaluated using 5-fold cross-validation for the 6-class classification problem. The error bars represent the standard deviation across folds.

To provide a detailed insight into the classification performance of the best-performing model, a single confusion matrix is presented as a representative example of the results across the five cross-validation folds. Specifically, Figure 6.35 reports the confusion matrix obtained from the third fold of the 5-fold cross-validation using the XGBoost model. While individual confusion matrices may exhibit slight variations across different folds, the reported matrix is indicative of the general classification trends observed throughout the evaluation process.

6.2.2 5-classes Classification Problem

6.2.3 Binary Classification Problem

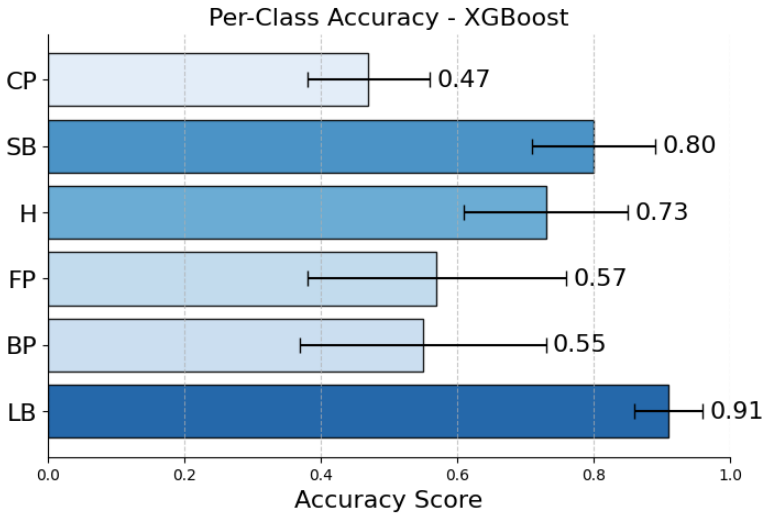


Figure 6.34: Per-class accuracy scores for XGBoost in the 6-class classification problem, evaluated using 5-fold cross-validation. The error bars represent the standard deviation across folds. The class labels correspond to the following smartphone positions: **LB** - Lower Back, **BP** - Back Pocket, **FP** - Front Pocket, **H** - Hand, **SB** - Shoulder Bag, and **CP** - Coat Pocket.

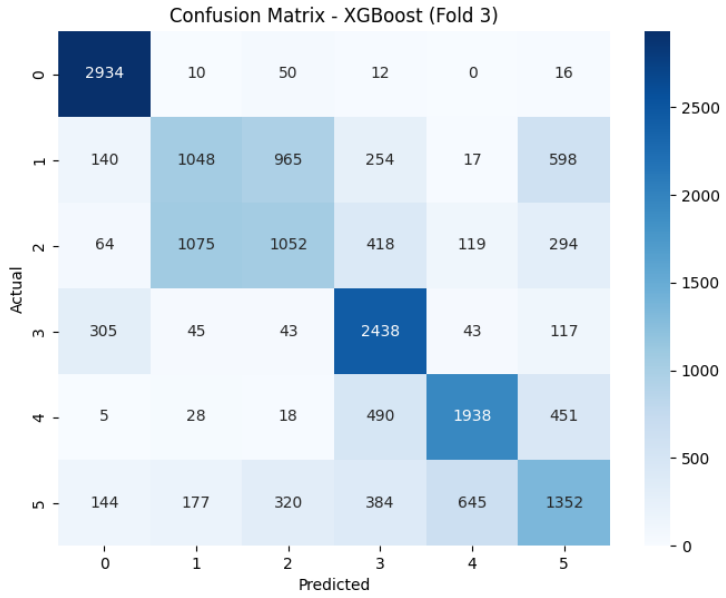


Figure 6.35: Confusion matrix for the XGBoost model, evaluated on the third fold of the 5-fold cross-validation. The rows represent the actual class labels, while the columns correspond to the predicted labels. The class mappings are as follows: **LB (0)** - Lower Back, **BP (1)** - Back Pocket, **FP (2)** - Front Pocket, **H (3)** - Hand, **SB (4)** - Shoulder Bag, and **CP (5)** - Coat Pocket. The matrix provides insights into the classification performance across different positions, highlighting correct predictions along the diagonal and misclassifications in the off-diagonal elements.

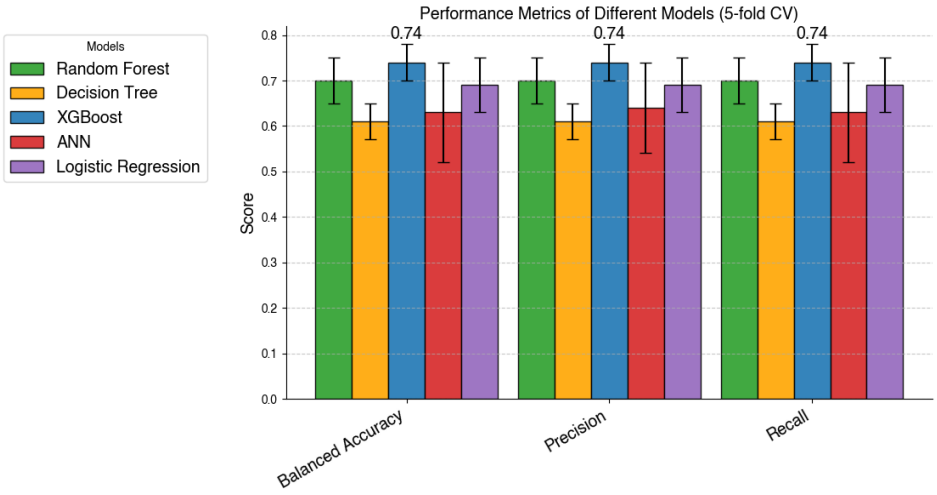


Figure 6.36: Performance metrics (Balanced Accuracy, Precision, Recall) for different classification models evaluated using 5-fold cross-validation for the 5-class classification problem. The error bars represent the standard deviation across folds.

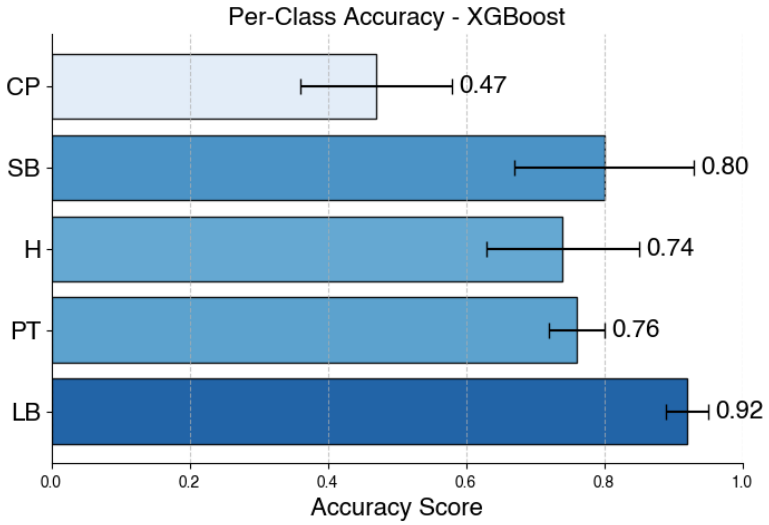


Figure 6.37: Per-class accuracy scores for XGBoost in the 5-class classification problem using 5-fold cross-validation. The error bars represent the standard deviation across folds. The class labels correspond to: **LB** - Lower Back, **PT** - Pocket Trousers (Back and Front merged), **H** - Hand, **SB** - Shoulder Bag, and **CP** - Coat Pocket.

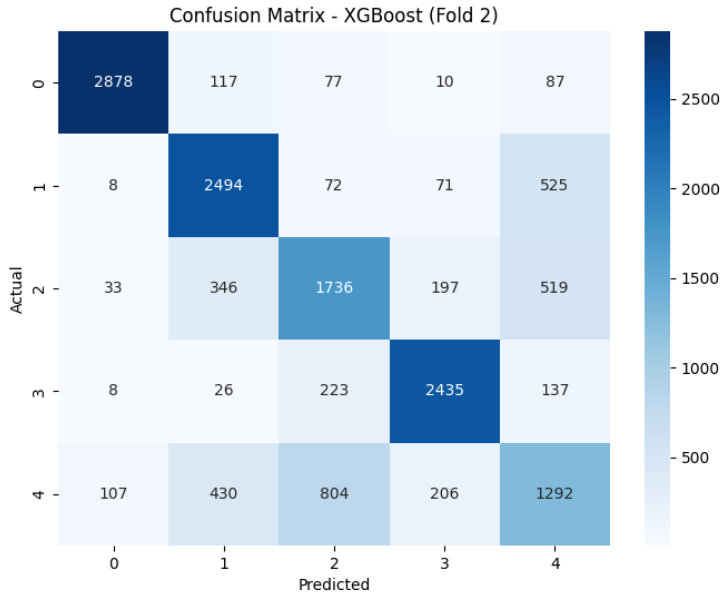


Figure 6.38: Confusion matrix for the XGBoost model in the 5-class classification problem, evaluated on Fold 2 of the 5-fold cross-validation. The rows represent the actual class labels, while the columns correspond to the predicted labels. The class mappings are as follows: **0** - Lower Back (LB), **1** - Pocket (Back and Front merged, PT), **2** - Hand (H), **3** - Shoulder Bag (SB), and **4** - Coat Pocket (CP).

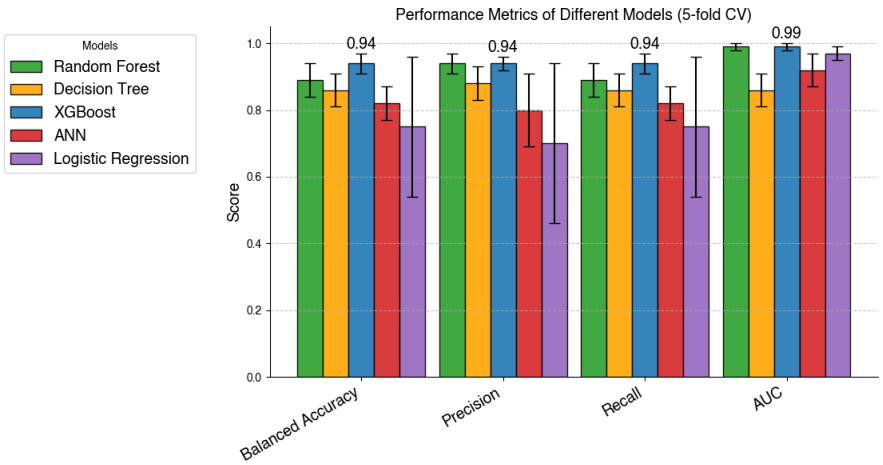


Figure 6.39: Performance metrics for different models in the binary classification problem (LB vs Other), evaluated using 5-fold cross-validation. The metrics reported include balanced accuracy, precision, recall, and AUC (Area Under the Curve). The error bars indicate the standard deviation across folds.

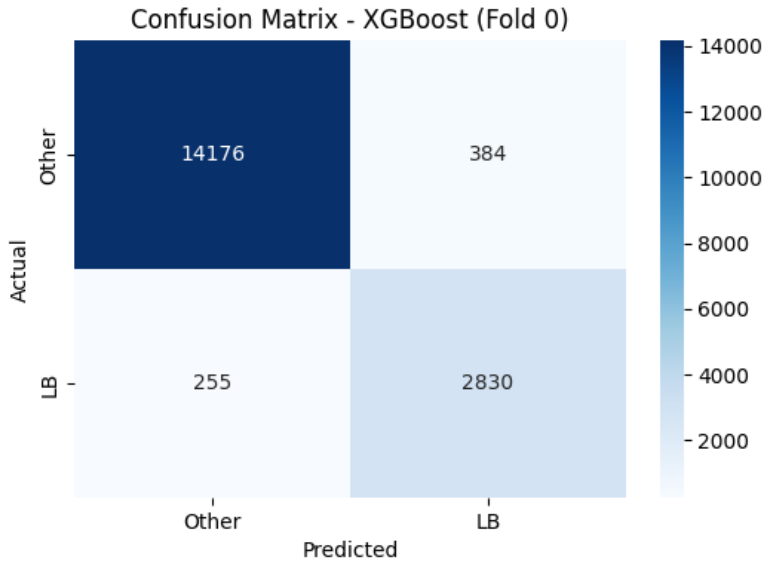


Figure 6.40: Confusion matrix for XGBoost in the binary classification problem (LB vs Other) using 5-fold cross-validation.

Chapter 7

Discussion

The overarching goal of this work was twofold. Initially, the aim was to validate the Mobilise-D pipeline by using inertial data collected from a standard smartphone, with its performance being compared against that of a dedicated research-grade device (an INDIP MIMU). Specifically, two complementary strategies were pursued:

1. Each of the four main modules of the pipeline—Gait Sequence Detection (GSD), Initial Contact Detection (ICD), Cadence (CAD), and Stride Length (SL)—was evaluated individually to determine its accuracy and robustness when applied to smartphone data.

2. The entire pipeline was applied end-to-end and its final outcome (walking speed) was assessed, thereby examining how smartphone-derived measurements aligned with those from the INDIP system in both in-lab and free-living scenarios.

By comparing a widely accessible consumer device (smartphone) with a specialized, research-grade IMU, it was sought to establish whether smartphones could serve as a practical alternative for reliable gait analysis.

The second objective of this study involved the development of a SLR model, considering that smartphone positioning varied significantly in free-living conditions.

In the following sections, the results of these investigations were presented and discussed, with a focus first on the individual modules, then on the full pipeline for walking speed estimation, and finally on the smartphone location recognition task.

7.1 Validation of the Mobilise-D Pipeline

7.1.1 GSD Module

The analysis of the smartphone data in both in-lab and free-living conditions revealed distinct distributional patterns for the detection metrics across the four algorithms ($\mathbf{GSD_A}$, $\mathbf{GSD_{A2}}$, $\mathbf{GSD_C}$, and $\mathbf{GSD_B}$). In the laboratory (see Figures 6.1, 6.2, 6.3, 6.4, and 6.5), the distributions of absolute error, accuracy, bias, precision, and recall were relatively narrow. In this controlled environment, $\mathbf{GSD_A}$—proposed as the optimal detector for a population of healthy adults—was observed to show competitive performance with high recall (i.e., nearly all true gait events were captured) and good accuracy. However, its precision was slightly lower compared to $\mathbf{GSD_{A2}}$, indicating that while $\mathbf{GSD_A}$ seldom missed a gait sequence, it tended to include a few more false positives. Meanwhile, both $\mathbf{GSD_C}$ and $\mathbf{GSD_B}$ exhibited lower median absolute errors and reduced variability; notably, $\mathbf{GSD_B}$ stood out for its minimal bias and well-balanced precision–recall trade-off.

In free-living conditions (refer to Figures 6.7, 6.8, 6.9, 6.10, and 6.11), the inherent variability of everyday movements led to broader distributions of errors and detection metrics for all algorithms. Although the accuracy remained high (typically in the 0.96–0.98 range) and the ICC values continued to be robust, the increased environmental variability caused the absolute errors and limits of agreement (LOA) to widen. In this scenario, the recall and bias distributions still favored $\mathbf{GSD_A}$, supporting its suitability for healthy adults; however, the improved precision and reduced bias observed for $\mathbf{GSD_B}$ suggested that a more balanced performance under real-world conditions might have been offered by the latter.

In summary, while $\mathbf{GSD_A}$ demonstrated excellent recall in the lab environment, a comprehensive evaluation of all metrics showed that $\mathbf{GSD_B}$ and $\mathbf{GSD_C}$ achieved lower absolute errors and tighter LOA in controlled settings. In free-living scenarios, despite an expected increase in variability, $\mathbf{GSD_A}$ remained competitive; nevertheless, $\mathbf{GSD_B}$ appeared to better manage the trade-offs between precision, bias, and recall. These findings underscored the necessity of considering multiple performance metrics simultaneously, as the optimal choice of detector might have varied depending on the specific application context.

Comparison of $\mathrm{GSD_A}$ Applied to Smartphone vs. IMU Data

Table 6.1 presented a detailed comparison of the performance of the $\mathrm{GSD_A}$ module when applied to data acquired from a smartphone and from the IMU of the INDIP kit, under both in-lab and free-living conditions.

In the laboratory setting, both systems exhibited very similar performance. The accuracy for both the smartphone and the IMU was approximately 0.87, and the recall was nearly 0.99, indicating that both devices were highly effective in

capturing the true gait events. Precision and specificity values were also comparable (around 0.74–0.75), suggesting that the proportion of false positives was minimal. The duration absolute error and bias, which provided insight into the consistency of the detected gait-sequence durations, were likewise very similar between the two devices (with the smartphone showing an absolute error of 3.86 s and a bias of 3.79 s, compared to 3.95 s and 3.65 s for the IMU, respectively). The ICC value of 0.99 in both cases further confirmed an excellent agreement with the reference measurements.

Under free-living conditions, despite the natural increase in variability due to unstructured movement, both devices continued to perform robustly. The accuracy rose to approximately 0.97, and specificity reached around 0.96 for both systems. Although the absolute error increased (to 4.78 s for the smartphone and 5.40 s for the IMU) and the bias values became slightly larger (3.22 s for the smartphone and 3.17 s for the IMU), the overall performance remained very similar, with the ICC still at 0.99. These metrics indicated that, even in more challenging real-world environments, the GSD_A module yielded consistent and reliable gait-sequence detection regardless of the acquisition device.

Overall, these results demonstrated that the smartphone, when paired with the GSD_A module, performed on par with the dedicated IMU of the INDIP kit.

7.1.2 ICD Module

The analysis of the smartphone data for initial contact detection in both in-lab and free-living conditions revealed distinct distributional patterns for the detection metrics across the three algorithms (ICD_A, ICD_D, and ICD_E). In the laboratory (see Figures 6.13, 6.14, 6.15, and 6.16), the overall distributions of absolute error, F1-score, precision, and recall were relatively narrow. In this controlled setting, $\mathbf{ICD_A}$—recommended for a healthy adult population—demonstrated excellent performance, exhibiting very low median absolute error and high recall, which indicated its outstanding ability to capture almost all true initial contacts. Although its precision was strong, the slight differences when compared to $\mathbf{ICD_D}$ suggested that while $\mathbf{ICD_A}$ rarely missed true events, it might occasionally have included some false detections. Meanwhile, $\mathbf{ICD_D}$ showed a comparable performance with a similar error spread but with a slightly more stable precision, whereas $\mathbf{ICD_E}$ exhibited a somewhat higher error distribution and lower median precision, pointing to a tendency for occasional misdetections.

When considering the free-living conditions (see Figures 6.17, 6.18, 6.19, and 6.20), the increased variability of real-world movements naturally led to broader distributions of errors and detection metrics across all algorithms. Although all methods showed a decrease in performance compared to the lab setting, the recall values remained high, ensuring that most true initial contacts were still detected.

Notably, $\mathbf{ICD_A}$ continued to perform strongly, maintaining high recall and balanced precision even in free-living scenarios. In contrast, $\mathbf{ICD_D}$ and $\mathbf{ICD_E}$ experienced a more noticeable spread in both absolute error and precision, indicating that their performance was more affected by the unstructured nature of free-living data.

In summary, while all three algorithms demonstrated robust initial contact detection under controlled conditions, the boxplots clearly showed that $\mathbf{ICD_A}$ offered the most favorable trade-off between low absolute error and high recall, making it particularly well-suited for applications involving healthy adults. Even though free-living conditions introduced greater variability, $\mathbf{ICD_A}$ remained competitive, supporting its use as the preferred method for initial contact detection in real-world scenarios.

Comparison of $\mathbf{ICD_A}$ Applied to Smartphone vs. INDIP Data

Table 6.2 presented a detailed comparison of the performance of the ICD_A module when applied to data acquired from a smartphone and from the MIMU of the INDIP kit, under both in-lab and free-living conditions.

In the laboratory setting, both systems exhibited very similar performance. The F1-scores were 0.904 for the smartphone and 0.906 for INDIP, indicating that both devices effectively balanced precision and recall in detecting initial contacts. The precision was also comparable (0.956 for the smartphone versus 0.958 for INDIP), while the recall values were nearly identical (0.860 vs. 0.862), demonstrating that the proportion of false positives and false negatives was minimal. Furthermore, the absolute error was extremely low at 0.069 s for both systems, with overlapping confidence intervals ([0.01, 0.18] s for the smartphone and [0.01, 0.17] s for INDIP). These findings confirmed that, under controlled conditions, the ICD_A module performed consistently regardless of the acquisition device.

Under free-living conditions, despite the increased variability inherent to real-world environments, the performance remained robust. The F1-scores were 0.842 for the smartphone and 0.838 for INDIP, and precision was 0.862 compared to 0.859, while recall values were 0.832 and 0.828, respectively. The absolute error remained similarly low, at 0.063 s for the smartphone and 0.065 s for INDIP, with both systems exhibiting comparable confidence intervals ([0.01, 0.17] s). These metrics suggested that even in less controlled, free-living scenarios, the ICD_A module yielded consistent and reliable detection of initial contacts across both platforms.

Overall, these results demonstrated that the smartphone, when paired with the ICD_A module, performed on par with the dedicated MIMU of the INDIP kit in both laboratory and free-living settings.

7.1.3 CAD Module

An examination of the three cadence-estimation algorithms ($\mathbf{CAD_A}$, $\mathbf{CAD_B}$, and $\mathbf{CAD_C}$) in both in-lab and free-living conditions revealed notable variations in their absolute and relative error distributions. In the controlled environment (see Figures 6.21 and 6.22), all three methods produced consistently accurate estimates of steps per minute, as indicated by relatively tight error spreads. However, $\mathbf{CAD_A}$ showed a somewhat broader range of outliers, suggesting that it was slightly more sensitive to minor fluctuations in step timing. By contrast, $\mathbf{CAD_C}$—the recommended algorithm—exhibited a particularly narrow distribution of relative error, indicating a strong alignment with reference values.

When shifting to free-living scenarios (see Figures 6.23 and 6.24), the inherent variability of daily activity became apparent, as the error distributions widened for all three algorithms. Despite this, $\mathbf{CAD_C}$ maintained comparatively lower median errors and a more contained spread, suggesting that it adapted effectively to irregular walking patterns. $\mathbf{CAD_B}$ also remained viable, though it appeared to yield more high-error outliers than $\mathbf{CAD_C}$, whereas $\mathbf{CAD_A}$ demonstrated larger overall deviations, potentially reflecting an increased sensitivity to real-world variability.

In summary, although each algorithm performed reliably under controlled conditions, $\mathbf{CAD_C}$ consistently showed the most stable performance in both laboratory and free-living environments. Its tighter error distributions and reduced bias highlighted its robustness for a wide range of gait analysis applications, whereas $\mathbf{CAD_A}$ and $\mathbf{CAD_B}$ appeared more susceptible to measurement fluctuations outside the lab.

Comparison of $\mathbf{CAD_C}$ Applied to Smartphone vs. INDIP Data

Table 6.3 presented a detailed comparison of the performance of the $\mathbf{CAD_C}$ algorithm when applied to cadence detection using data acquired from a smartphone and from the INDIP MIMU, in both in-lab and free-living conditions.

In the laboratory setting, the absolute error in cadence estimation was very similar between the two systems. The smartphone yielded an average absolute error of 1.86 steps/min (with a confidence interval of [0.07, 5.70]), while the INDIP system showed an error of 1.75 steps/min ([0.05, 5.23]). Likewise, the relative error was nearly identical, with the smartphone recording 0.00% and the INDIP 0.01%, both within overlapping confidence intervals. These results indicated that, under controlled conditions, the $\mathbf{CAD_C}$ algorithm performed comparably regardless of whether the data was sourced from a smartphone or a dedicated MIMU.

Under free-living conditions, as expected, the variability increased; however, the performance remained essentially equivalent between the two acquisition methods. The absolute error was 5.66 steps/min for the smartphone and 5.77 steps/min for

the INDIP, with overlapping confidence intervals ([0.28, 21.08] vs. [0.28, 21.96]). Similarly, the relative error was consistently around 0.01% for both devices. These findings confirmed that the **CAD_C** module maintained robust and reliable cadence detection even in real-world scenarios, independent of the sensor used.

Overall, the near-identical performance of **CAD_C** across both in-lab and free-living environments underscored its suitability for accurate cadence estimation using either smartphone or INDIP MIMU data.

7.1.4 SL Module

The analysis of the smartphone data for stride-length estimation in both in-lab and free-living conditions revealed distinct distributional patterns for the estimation metrics across the two algorithms (SL_A and SL_B). In the laboratory (see Figures 6.25, 6.26, and 6.27), the overall distributions of absolute error, bias, and relative error were relatively narrow. In this controlled setting, **SL_B**—recommended for a healthy adult population—demonstrated outstanding performance, exhibiting a very low median absolute error and minimal bias, which indicated strong consistency with the reference stride lengths. Although SL_A achieved acceptable results, its error distributions and bias were slightly broader, suggesting occasional over- or underestimation.

Under free-living conditions (see Figures 6.28, 6.29, and 6.30), the increased variability inherent in real-world movements naturally led to broader error distributions for both algorithms. Nonetheless, **SL_B** maintained a comparatively lower median absolute error and a tighter bias distribution, reflecting its robust performance even in unstructured environments. In contrast, SL_A showed a larger spread in errors, indicating a higher sensitivity to irregular gait patterns during daily activities.

In summary, while both algorithms performed adequately in controlled settings, the boxplots clearly indicated that **SL_B** offered the most favorable trade-off between low absolute error and minimal bias. This made it the preferred approach for stride-length estimation in both laboratory and free-living scenarios.

Comparison of SL_B Applied to Smartphone vs. INDIP Data

Table 6.4 presented a detailed comparison of the performance of the SL_B module for stride length estimation when applied to data acquired from a smartphone and from the INDIP MIMU, under both in-lab and free-living conditions.

In the laboratory setting, both systems exhibited nearly identical performance. The absolute error was 0.17 m for the smartphone and 0.18 m for INDIP, with overlapping confidence intervals ([0.03, 0.33] m and [0.03, 0.35] m, respectively). Likewise, the relative error was consistent between the two, with values of -0.13% for both modalities, and the bias was the same at -0.18 m. These metrics indicated

that in a controlled environment, the SL_B module yielded equivalent and reliable stride length estimates regardless of the acquisition device.

Under free-living conditions, despite the naturally increased variability, the performance remained robust and comparable between the smartphone and INDIP. The absolute error was 0.11 m for the smartphone and 0.12 m for INDIP, while the relative error was nearly identical (0.02% versus 0.03%). The bias was also consistent at -0.02 m for both systems, with similar confidence intervals. These results confirmed that the SL_B module maintained its accuracy and reliability in real-world settings, independent of the sensor used.

Overall, these findings demonstrated that the smartphone, when paired with the SL_B module, provided stride length estimates that were on par with those obtained from the dedicated INDIP MIMU, both in laboratory and free-living conditions.

7.1.5 Full Pipeline

Table 6.5 summarized the walking-speed results obtained by applying the full processing pipeline to data from both the smartphone and the INDIP IMU, under in-lab and free-living conditions. In the laboratory (upper panel), the absolute errors for the smartphone and IMU were 0.15 m/s^2 and 0.16 m/s^2, respectively, with largely overlapping confidence intervals. The relative error and bias values were similarly close (both around -0.15), indicating that the smartphone performed on par with the dedicated IMU in estimating walking speed in a controlled environment.

Under free-living conditions (lower panel), the absolute error decreased to around 0.09 m/s^2 for both devices, highlighting that even with the increased variability of real-world movement, the pipeline maintained robust performance. The relative error remained near zero, and the bias was approximately -0.04 m/s^2 for both the smartphone and IMU. These findings suggested that the system effectively adapted to unstructured daily activities, yielding comparable results regardless of the acquisition modality.

Figures **??** and **??** presented Bland-Altman plots of median walking speed per subject for in-lab and free-living data, respectively. In both plots, the mean difference (solid gray line) lay very close to zero, and the limits of agreement (dashed red and blue lines) remained relatively narrow. This further confirmed the strong agreement between the smartphone and IMU measurements, with minimal systematic offsets across different walking speeds.

Overall, these results demonstrated that the full pipeline, when applied to smartphone data, yielded walking-speed estimates that closely matched those obtained from a research grade device like the INDIP IMU. The small discrepancies observed—both in controlled and free-living conditions—were well within acceptable bounds, reinforcing the viability of using smartphones for accurate gait assessment in various settings.

7.2 Smartphone Location Recognition Model

7.2.1 6-Classes Classification Problem

Figure 6.33 shows the performance metrics (balanced accuracy, precision, and recall) for five different classification models, all evaluated via 5-fold cross-validation on the 6-class smartphone-location recognition task. The results indicate a generally consistent level of performance across models, with XGBoost slightly outperforming the others. Notably, balanced accuracy, precision, and recall scores cluster around the 0.65–0.70 range, reflecting a reasonably balanced capability to detect each class without heavily favoring one over the others.

A closer examination of the per-class accuracy for the XGBoost model (Figure 6.34) reveals that some classes, such as **LB** (Lower Back), achieve high accuracy (above 0.90), while others, notably **CP** (Coat Pocket), are more challenging, with an accuracy around 0.47. Intermediate performances are observed for positions like **SB** (Shoulder Bag) and **H** (Hand), suggesting that although the model can reliably recognize certain placements, it struggles with others due to less distinctive features.

Furthermore, the confusion matrix shown in Figure 6.35 highlights a significant issue between the **BP** (Back Pocket) and **FP** (Front Pocket) classes. Both classes exhibit low accuracy and high standard deviation, which indicates that the features extracted from these positions are not sufficiently discriminative. This outcome suggests that the two placements, being located on the same anatomical region, generate similar inertial signals. Moreover, when considering future developments in gait analysis algorithms that may utilize data from these placements, it might not be necessary to differentiate between front and back pockets if the signals are essentially equivalent. For these reasons, a decision was made to merge **BP** and **FP** into a single *Trousers Pocket* (**PT**) class. The impact of this merging on the overall model performance will be discussed in the subsequent section.

In summary, while the 6-class classification problem demonstrates the model's overall capability to recognize multiple smartphone positions with moderate to high accuracy, the analysis indicates that certain classes—specifically **BP** and **FP**—pose a challenge due to insufficiently distinctive features. Merging these into a single class appears to be a justified strategy both from a current performance perspective and for the development of future gait analysis algorithms.

7.2.2 5-Classes Classification Problem

Following the merging of the Back Pocket (BP) and Front Pocket (FP) classes into a single **Pocket Trousers (PT)** category, the classification task is reduced to five classes: **LB** (Lower Back), **PT** (Pocket Trousers), **H** (Hand), **SB** (Shoulder

Bag), and **CP** (Coat Pocket). Figure 6.36 illustrates the performance metrics (balanced accuracy, precision, and recall) for the same five models (Random Forest, Decision Tree, XGBoost, ANN, and Logistic Regression), each evaluated with 5-fold cross-validation. Notably, the average scores for all three metrics cluster around 0.74, suggesting a slight overall improvement compared to the 6-class scenario.

Focusing on the XGBoost model, the best performing model, Figure 6.37 provides a breakdown of per-class accuracy. **LB** maintains the highest accuracy (approximately 0.92), while **CP** remains the most challenging class, with an accuracy near 0.47. **PT** achieves about 0.76, indicating that merging the two pocket classes alleviated some of the confusion observed in the 6-class problem, where BP and FP had notably low accuracies and high variability. **SB** (Shoulder Bag) and **H** (Hand) both show moderate accuracy levels (around 0.80 and 0.74, respectively), confirming that these positions are reasonably distinct yet still prone to occasional misclassification.

The confusion matrix for the XGBoost model on Fold 2 (Figure 6.38) further illustrates these trends. Correct predictions are concentrated on the main diagonal, with **LB** showing very few misclassifications, while **CP** exhibits a higher degree of confusion with other classes. Notably, **PT** displays a reduced rate of misclassifications compared to the separate **BP** and **FP** classes, reinforcing the decision to merge them into a single category.

In summary, the 5-class classification problem demonstrates that merging the pocket-related classes yields a modest yet tangible improvement in overall model performance. By reducing the overlap in sensor signals associated with front and back pockets, the classification task becomes more tractable, resulting in balanced accuracy, precision, and recall scores of around 0.74 for the best-performing methods. Although some classes (e.g., **CP**) remain challenging, these findings suggest that consolidating similar positions can enhance both model stability and interpretability in smartphone location recognition.

7.2.3 Binary Classification Problem: LB vs Other

Finally, a binary classification task was carried out to distinguish the **LB** placement from all other positions. Although the lower back is not a common spot for everyday smartphone use, it is the standard location for many gait-analysis protocols, including Mobilise-D. Demonstrating robust recognition of **LB** is thus valuable, as it enables the integration of a location-recognition step into a broader smartphone-based gait-analysis framework.

Figure 6.39 presents the balanced accuracy, precision, recall, and AUC (Area Under the Curve) for five different models, all evaluated via 5-fold cross-validation. The results are notably high, with XGBoost, for example, surpassing 0.90 in each metric and reaching 0.99 in AUC. These findings confirm that distinguishing **LB**

from other placements is considerably more straightforward than the more granular multi-class tasks.

The confusion matrix for XGBoost (Figure 6.40) further supports this conclusion, showing a large number of correctly identified **Other** instances (14,176) and a similarly high count of correctly classified **LB** samples (2,830). Misclassifications remain minimal, indicating that the model reliably detects the **LB** position with little confusion. From a practical standpoint, this high level of accuracy suggests that the location-recognition model can be readily incorporated into smartphone-based gait-analysis pipelines. In other words, once the lower-back position is confidently identified, specialized algorithms (e.g., those employed by Mobilise-D) can be selectively applied to the sensor data, ensuring accurate gait metrics when the smartphone is placed at **LB**.

Chapter 8

Conclusions

8.1 General results

This thesis aimed to validate the Mobilise-D pipeline for gait analysis while also developing a classification model for the task of smartphone location recognition. Data acquired from both controlled (in-lab) and real-world (free-living) environments—using an Android smartphone and a research-grade device—were utilized for both pipeline validation and model training. In particular, the INDIP sensor kit was employed as the reference system, providing a benchmark to compare performance metrics during pipeline validation and serving to extract walking sequences from all recordings for training the location recognition model.

Beyond validating existing methodologies, this work sought to enhance the generalizability of gait analysis algorithms to accommodate real-world smartphone usage. While traditional IMU-based approaches rely on standardized placements, smartphones introduce variability in positioning that must be accounted for to ensure reliable movement assessment. The development of classification models capable of identifying smartphone placement is therefore a crucial step toward optimizing data interpretation and refining digital mobility assessment tools.

The output of this work were the validation results, which highlighted that the smartphone—a widely available and accessible device—represented a viable hardware alternative to dedicated research-grade sensors. Moreover, the construction of the classification model for smartphone location recognition demonstrated that the model reliably distinguished between different device placements across various classification tasks (multi-class, 5-class, and binary), with performance metrics consistently improving after merging classes with similar inertial signatures. In particular, the results for the smartphone location recognition model were robust and excellent for identifying the LB (Lower Back) position, although performance for the other placements was slightly lower.

Overall, the findings confirmed that smartphone-based measurements could be effectively integrated within the Mobilise-D analytical framework, delivering comparable accuracy to that of the reference system for gait parameters while also enabling reliable device location recognition. These results underscored the potential of using smartphones as an accessible alternative to research-grade sensors in gait analysis applications. However, limitations such as variability in certain device placements and the need for more refined, position-aware analytical methods were also revealed. As such, while this study provided a strong foundation, it simultaneously pointed to several avenues for future research. In the following section, the study's limitations and potential developments will be discussed, outlining strategies to further enhance the robustness and generalizability of smartphone-based gait analysis.

8.2 Future directions

Validation of smartphone-acquired data using the Mobilise-D pipeline across different cohorts

One limitation of this study, particularly regarding the validation of the Mobilise-D pipeline, is that although data were acquired from a substantial number of participants, they were exclusively collected from healthy adults. Consequently, the current findings may not fully generalize to populations with irregular gait patterns or more complex conditions. Future research should extend the application of this pipeline to smartphone data gathered from other cohorts—similar to the approach used in the European Mobilise-D Technical Validation Study [45]—to evaluate whether smartphones remain a reliable tool for gait analysis in populations with conditions such as Parkinson's disease, musculoskeletal disorders, or post-stroke impairments.

Current constraints and future improvements for Smartphone Location Recognition

Despite the optimal performance in correctly classifying the LB placement—which supported the model's applicability for recognizing this specific position—it was acknowledged that the lower back is not a commonly used smartphone location. In contrast, the results obtained for other, more frequently used placements indicated that further improvements were needed to enhance the model's accuracy in recognizing these positions. Future work should explore why the current features did not provide sufficiently discriminative information for differentiating among the various placements, particularly for the coat pocket placement, which remains an open challenge.

Another limitation is that the model was trained exclusively on walking data. Although this design choice was made to better extract distinctive features from the signals corresponding to different placements, it meant that the model never encountered signals representing transitions between placements—as is common in typical smartphone usage. Incorporating these transitional signals would be crucial for continuous monitoring, allowing the model to identify transition segments and exclude them from the gait parameter extraction process, since they do not provide the data that the analytical pipelines expect.

Bibliography

[1] Reed Ferber, Sean T. Osis, Jennifer L. Hicks, and Scott L. Delp. «Gait Biomechanics in the Era of Data Science». In: *Journal of Biomechanics* (2016). DOI: 10.1016/j.jbiomech.2016.10.033. URL: http://dx.doi.org/10.1016/j.jbiomech.2016.10.033 (cit. on pp. 1, 11).

[2] A. Muro-de-la-Herran, B. García-Zapirain, and A. Méndez-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. URL: https://doi.org/10.3390/s140203362 (cit. on p. 1).

[3] R. Baker. «Gait analysis methods in rehabilitation». In: *Journal of NeuroEngineering and Rehabilitation* 3.1 (2006), p. 4. DOI: 10.1186/1743-0003-3-4. URL: https://doi.org/10.1186/1743-0003-3-4 (cit. on p. 1).

[4] E. Giannouli, O. Bock, S. Mellone, and W. Zijlstra. «Mobility in old age: Capacity is not performance». In: *BioMed Research International* 2016 (2016), p. 3261567. DOI: 10.1155/2016/3261567. URL: https://doi.org/10.1155/2016/3261567 (cit. on pp. 1, 17).

[5] M. A. Khan, S. Lee, H. Park, and M. Lee. «A wearable sensor-based gait analysis framework for rehabilitation». In: *Sensors* 15.8 (2015), pp. 20795–20821. DOI: 10.3390/s150820795. URL: https://doi.org/10.3390/s150820795 (cit. on p. 2).

[6] Z. Zhang, L. Wang, and J. Wang. «Smartphone-based gait analysis: Current methodologies and future perspectives». In: *IEEE Sensors Journal* 21.14 (2021), pp. 15980–15991. DOI: 10.1109/JSEN.2021.3075198. URL: https://doi.org/10.1109/JSEN.2021.3075198 (cit. on p. 2).

[7] Akram Bayat, Marc Pomplun, and Duc A. Tran. «A Study on Human Activity Recognition Using Accelerometer Data from Smartphones». In: *Procedia Computer Science*. Vol. 34. 2014, pp. 450–457. DOI: 10.1016/j.procs.2014.07.009. URL: https://doi.org/10.1016/j.procs.2014.07.009 (cit. on pp. 2, 16).

[8] Sharon Olsen, Usman Rashid, David Barbado, Priyadharshini Suresh, Gemma Alder, Imran Khan Niazi, and Denise Taylor. «The validity of smartphone-based spatiotemporal gait measurements during walking with and without head turns: Comparison with the GAITRite® system». In: *Journal of Biomechanics* 162 (2024). DOI: 10.1016/j.jbiomech.2023.111899. URL: https://doi.org/10.1016/j.jbiomech.2023.111899 (cit. on pp. 2, 16).

[9] C. Kirk, A. Küderle, M. E. Micó-Amigo, et al. «Mobilise-D insights to estimate real-world walking speed in multiple conditions with a wearable device». In: *Scientific Reports* 14 (2024), p. 1754. DOI: 10.1038/s41598-024-51766-5. URL: https://doi.org/10.1038/s41598-024-51766-5 (cit. on pp. 2, 61–64, 68).

[10] MobGap development team. *MobGap Documentation*. 2025. URL: https://mobgap.readthedocs.io/en/stable/ (cit. on pp. 2, 61).

[11] M. Encarna Micó-Amigo, Tecla Bonci, Anisoara Paraschiv-Ionescu, et al. «Assessing real-world gait with digital technology? Validation, insights and recommendations from the Mobilise-D consortium». In: *Journal of Neuro-Engineering and Rehabilitation* 20 (2023), p. 78. DOI: 10.1186/s12984-023-01198-5. URL: https://doi.org/10.1186/s12984-023-01198-5 (cit. on pp. 2, 61–63, 68, 79).

[12] Jacquelin Perry. *Gait Analysis: Normal and Pathological Function*. Slack Incorporated, 1992 (cit. on pp. 6–10).

[13] Thurmon E. Lockhart and Jeawon Kim. *Biomechanics of Walking and Running*. Springer, 2013 (cit. on p. 7).

[14] Chayanika Dhiman. «Gait Analysis: A Review». In: *International Journal of Computer Applications* 29.6 (2011), pp. 38–45 (cit. on pp. 8, 9).

[15] J. Perry. «Clinical gait analyzer». In: *Bulletin of Prosthetics Research* (1974) (cit. on p. 10).

[16] D. Sutherland and J. Hagy. «Measurement of gait movements from motion picture film». In: *British Journal of Bone and Joint Surgery* (1972) (cit. on p. 10).

[17] Richard Baker. «The history of gait analysis before the advent of modern computers». In: *Gait & Posture* 26 (2007), pp. 331–342. DOI: 10.1016/j.gaitpost.2006.10.014. URL: https://doi.org/10.1016/j.gaitpost.2006.10.014 (cit. on p. 10).

[18] R. Sheldon. «Biomechanics of Gait Analysis». In: *Journal of Biomechanics* 37.4 (2004), pp. 543–550 (cit. on p. 10).

[19] Weijun Tao, Tao Liu, Rencheng Zheng, and Hutian Feng. «Gait Analysis Using Wearable Sensors». In: *Sensors* 12.2 (2012). DOI: 10.3390/s120202255. URL: https://doi.org/10.3390/s120202255 (cit. on pp. 11, 13).

[20] Michael W. Whittle. *Gait Analysis: An Introduction*. Butterworth-Heinemann, 1996 (cit. on pp. 11–13).

[21] Meg E. Morris, Robert Iansek, T. Alexander Matyas, and Jeffrey J. Summers. «Abnormalities in the stride length-cadence relation in Parkinsonian gait». In: *Movement Disorders* 13.1 (1998), pp. 61–69 (cit. on p. 11).

[22] Francesco Temporiti and Roberto Gatti. «Gait analysis in patients after bilateral versus unilateral total hip arthroplasty». In: *Gait & Posture* 70 (2019), pp. 74–79 (cit. on p. 11).

[23] Tomas Oberg, Anna Karsznia, and Katarina Oberg. «Basic gait parameters: Reference data for normal subjects, 10-79 years of age». In: *Journal of Rehabilitation Research and Development* 30.2 (1993), pp. 210–223 (cit. on pp. 12, 13).

[24] Elisa Digo. «Comparison of different inertial sensors setups and algorithms to estimate gait spatio-temporal parameters». MA thesis. Politecnico di Torino, 2018 (cit. on p. 13).

[25] Alana de Mello Souza and Marcelo R. Stemmer. «Extraction and classification of human gait parameters using Kinect sensor». In: *Journal of Control, Automation and Electrical Systems* 29 (2018), pp. 586–604 (cit. on p. 13).

[26] Wiebren Zijlstra and At L. Hof. «Assessment of spatio-temporal gait parameters from trunk accelerations during human walking». In: *Gait & Posture* 18.2 (2003), pp. 1–10 (cit. on pp. 14, 66).

[27] Wiebren Zijlstra and Astrid Zijlstra. «Assessment of gait parameters in older persons: A comparison of accelerometry and clinical observation». In: *Gait & Posture* 19.1 (2004), pp. 1–7 (cit. on pp. 14, 17).

[28] Ugo Della Croce and Maurizio Mancini. «Application of Inertial Measurement Units in Clinical Gait Analysis: Current Trends and Future Directions». In: *Handbook of Human Motion*. Ed. by B. Muller and S. Wolf. Springer, 2018, pp. 1–20 (cit. on p. 14).

[29] S. Bovonsunthonchai, R. Vachalathiti, A. Pisarnpong, F. Khobhun, and V. Hiengkaew. «Spatiotemporal gait parameters for patients with Parkinson's disease compared with normal individuals». In: *Physiotherapy Research International* 19.3 (Sept. 2014), pp. 158–165. DOI: 10.1002/pri.1579. URL: https://doi.org/10.1002/pri.1579 (cit. on pp. 15, 24).

[30] Daniele Trojaniello, Andrea Cereatti, and Ugo Della Croce. «Assessment of spatio-temporal parameters of gait in stroke patients using wearable inertial sensors». In: *Gait & Posture* 41.3 (2015), pp. 805–810 (cit. on pp. 15, 25).

[31] R. A. W. Felius, M. Geerars, S. M. Bruijn, J. H. van Dieën, N. C. Wouda, and M. Punt. «Reliability of IMU-Based Gait Assessment in Clinical Stroke Rehabilitation». In: *Sensors (Basel)* 22.3 (Jan. 2022), p. 908. DOI: 10.3390/s22030908. URL: https://doi.org/10.3390/s22030908 (cit. on p. 15).

[32] S. Baklouti, A. Chaker, T. Rezgui, A. Sahbani, S. Bennour, and M. A. Laribi. «A Novel IMU-Based System for Work-Related Musculoskeletal Disorders Risk Assessment». In: *Sensors* 24 (2024), p. 3419. DOI: 10.3390/s24113419. URL: https://doi.org/10.3390/s24113419 (cit. on p. 15).

[33] Sangrok Kang, Soo-Mi Choi, Sang-Rae Lee, Ho-Joon Lee, Kyung-Yeon Kim, and Young-Seok Kim. «A novel walking detection and step counting algorithm using unconstrained smartphones». In: *Sensors* 18.12 (2018), p. 4265 (cit. on p. 16).

[34] Pilar Serra-Añó, Mónica Balasch-Bernat, Laura López-Bueno, Pedro Pérez-Soriano, and Xavier García-Massó. «Mobility assessment in people with Alzheimer's disease using smartphone sensors: A feasibility study». In: *Sensors* 19.23 (2019), p. 5153 (cit. on p. 16).

[35] GAITRite. *GAITRite - The Gold Standard in Gait Analysis*. Accessed: 2024-02-05. 2024. URL: https://www.gaitrite.com/ (cit. on p. 16).

[36] Jannis Hannink, Tim Kautz, Cristian F. Pasluosta, Bernhard Gasser, Jochen Klucken, and Bjoern M. Eskofier. «Mobile stride length estimation with deep convolutional neural networks». In: *IEEE Journal of Biomedical and Health Informatics* 21.3 (2017), pp. 568–576. DOI: 10.1109/JBHI.2016.2633772 (cit. on p. 16).

[37] Ilya Galperin, Isaac Hillel, Silvia Del Din, Edith M.J. Bekkers, Alice Nieuwboer, Lynn Rochester, and Jeffrey M. Hausdorff. «Associations between daily-living physical activity and mobility in patients with Parkinson's disease». In: *Parkinsonism & Related Disorders* 62 (2019), pp. 85–90. DOI: 10.1016/j.parkreldis.2019.01.029 (cit. on p. 17).

[38] Vectornav. *Inertial Navigation Primer*. 2022. URL: https://www.vectornav.com/resources/inertial-navigation-primer/theory-of-operation/theory-inertial (cit. on p. 18).

[39] Zakriya Mohammed, Ibrahim M. Elfadel, and Mahmoud Rasras. «Monolithic Multi Degree of Freedom (MDoF) Capacitive MEMS Accelerometers». In: *Micromachines* 9 (2018). DOI: 10.3390/mi9010009. URL: https://doi.org/10.3390/mi9010009 (cit. on pp. 19–22).

[40] Riccardo Antonello and Roberto Oboe. «MEMS Gyroscopes for Consumers and Industrial Applications». In: InTech, 2011. ISBN: 978-953-307-170-1. DOI: 10.5772/17689. URL: https://doi.org/10.5772/17689 (cit. on p. 21).

[41] J. Bernstein, S. Cho, A. T. King, A. Kourepinis, P. Maciel, and M. Weinberg. «Micromachined comb-drive tuning fork rate gyroscope». In: *IEEE Micro Electro Mechanical Systems*. IEEE, 1993, pp. 143–148. DOI: 10.1109/memsys.1993.296932. URL: https://doi.org/10.1109/memsys.1993.296932 (cit. on p. 22).

[42] W. A. Gill, I. Howard, I. Mazhar, and K. McKee. «A Review of MEMS Vibrating Gyroscopes and Their Reliability Issues in Harsh Environments». In: *Sensors* 22.19 (2022). DOI: 10.3390/s22197405. URL: https://doi.org/10.3390/s22197405 (cit. on p. 22).

[43] Sebastian O. H. Madgwick, Andrew J. L. Harrison, and Ravi Vaidyanathan. «Estimation of IMU and MARG orientation using a gradient descent algorithm». In: *2011 IEEE International Conference on Rehabilitation Robotics*. 2011, pp. 1–7. DOI: 10.1109/ICORR.2011.5975346 (cit. on p. 23).

[44] Felix Kluge and Claudia Mazzà. «Towards a consensus on the use of wearable technology in gait analysis: A framework for the digital monitoring of mobility». In: *PLOS One* 16.3 (2021), e0247645 (cit. on pp. 23, 25).

[45] C. Mazzà, L. Alcock, K. Aminian, C. Becker, S. Bertuletti, T. Bonci, 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. URL: https://doi.org/10.1136/bmjopen-2021-050785 (cit. on pp. 23, 25, 61, 123).

[46] Stefano Bertuletti, Andrea Cereatti, Daniele Comotti, Michele Caldara, and Ugo Della Croce. «Static and Dynamic Accuracy of an Innovative Miniaturized Wearable Platform for Short Range Distance Measurements for Human Movement Applications». In: *Sensors* 17.7 (2017). ISSN: 1424-8220. DOI: 10.3390/s17071492. URL: https://www.mdpi.com/1424-8220/17/7/1492 (cit. on pp. 24, 25, 31).

[47] Francesca Salis and Andrea Cereatti. «A method for detecting gait events using low-resolution pressure insoles». In: *Journal of Biomechanics* 47.6 (2014), pp. 1585–1588 (cit. on pp. 24, 25).

[48] Francesca Salis and Andrea Cereatti. «INDIP: A multi-sensor wearable system for real-world gait analysis». In: *SIAMOC Conference* (2019) (cit. on pp. 24, 25).

[49] Francesca Salis and Andrea Cereatti. «Technical validation of a multi-sensor wearable system (INDIP) for real-world gait analysis». In: *Frontiers in Bioengineering and Biotechnology* 11 (2023), p. 1145 (cit. on pp. 24, 25, 31, 69).

[50] Luca Palmerini and Lorenzo Chiari. «Standardization of wearable data collection and processing: The Mobilise-D project». In: *Scientific Data* 10 (2023), p. 245 (cit. on pp. 24, 48).

[51] L. Brognara. «Gait Assessment Using Smartphone Applications in Older Adults: A Scoping Review». In: *Geriatrics (Basel)* 9.4 (July 2024), p. 95. DOI: 10.3390/geriatrics9040095. URL: https://doi.org/10.3390/geriatrics9040095 (cit. on p. 27).

[52] Daniel Klein and Timo Becker. «Enhancing pedestrian dead reckoning through smartphone position recognition». In: *Sensors* 20.18 (2020), p. 5034 (cit. on pp. 27, 28).

[53] Jing Zhou, Xinyu Wang, and Zhen Liu. «Deep learning-based human activity recognition using smartphone sensors». In: *IEEE Transactions on Neural Networks and Learning Systems* 30.3 (2019), pp. 818–827 (cit. on p. 28).

[54] Marc Daniel, Tobias Klein, and Soo-Mi Kim. «Deep learning-based smartphone position recognition for robust activity recognition». In: *Pattern Recognition Letters* 146 (2021), pp. 72–79 (cit. on p. 28).

[55] F. Salis, S. Bertuletti, K. Scott, M. Caruso, T. Bonci, E. Buckley, U. Della Croce, C. Mazzà, and A. Cereatti. «A wearable multi-sensor system for real world gait analysis». In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019 (cit. on p. 32).

[56] M. Caruso, A. M. Sabatini, D. Laidig, T. Seel, M. Knaflitz, U. Della Croce, and A. Cereatti. «Analysis of the Accuracy of Ten Algorithms for Orientation Estimation Using Inertial and Magnetic Sensing under Optimal Conditions: One Size Does Not Fit All». In: *Sensors* 21 (2021), p. 2543. DOI: 10.3390/s21072543. URL: https://doi.org/10.3390/s21072543 (cit. on p. 35).

[57] A. Küderle. *Mobilise-D Technical Validation Study (TVS) dataset*. 2024. DOI: 10.5281/zenodo.13899385. URL: http://doi.org/10.5281/zenodo.13899385 (cit. on p. 62).

[58] Tal Iluz, Eran Gazit, Talia Herman, Eliot Sprecher, Marina Brozgol, Nir Giladi, Anat Mirelman, and Jeffrey M. Hausdorff. «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 (2014), p. 48. DOI: 10.1186/1743-0003-11-48. URL: http://www.jneuroengrehab.com/content/11/1/48 (cit. on p. 65).

[59] MobGap Developers. *MobGap: Gait Sequence Detection (GsdIluz)*. Online documentation. Accessed: 2024-02-20. 2024. URL: https://mobgap.readthedocs.io/en/stable/modules/generated/gait_sequences/mobgap.gait_sequences.GsdIluz.html (cit. on p. 65).

[60] John McCamley, Marcello Donati, Emiliano Grimpampi, and Claudia Mazza. «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.2012.03.019. URL: https://doi.org/10.1016/j.gaitpost.2012.03.019 (cit. on p. 65).

[61] Anisoara Paraschiv-Ionescu and Kamiar Aminian. «Real-world ambulatory monitoring: New frontiers in assessment and rehabilitation of mobility». In: *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 4226–4229. DOI: 10.1109/EMBC44109.2020.9175662. URL: https://doi.org/10.1109/EMBC44109.2020.9175662 (cit. on p. 65).

[62] MobGap Developers. *MobGap: Initial Contact Detection (IcdIonescu)*. Online documentation. Accessed: 2024-02-20. 2024. URL: https://mobgap.readthedocs.io/en/stable/modules/generated/initial_contacts/mobgap.initial_contacts.IcdIonescu.html (cit. on p. 66).

[63] MobGap Developers. *MobGap: Cadence Estimation (CadFromIc)*. Online documentation. Accessed: 2024-02-20. 2024. URL: https://mobgap.readthedocs.io/en/stable/modules/generated/cadence/mobgap.cadence.CadFromIc.html (cit. on p. 66).

[64] Amin Soltani, Anisoara Paraschiv-Ionescu, Hamid Dejnabadi, and Kamiar Aminian. «Stride length estimation from a single trunk IMU: A machine learning versus biomechanical modeling approach». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021), pp. 1116–1125. DOI: 10.1109/TNSRE.2021.3084091. URL: https://doi.org/10.1109/TNSRE.2021.3084091 (cit. on p. 66).

[65] MobGap Developers. *MobGap: Stride Length Estimation (SlZijlstra)*. Online documentation. Accessed: 2024-02-20. 2024. URL: https://mobgap.readthedocs.io/en/stable/modules/generated/stride_length/mobgap.stride_length.SlZijlstra.html (cit. on p. 67).