



**Politecnico
di Torino**

Politecnico di Torino

Master's Degree in Biomedical Engineering

Graduation Session July 2026

Context-Aware Assessment of Real-World Locomotion Using Head-Worn Sensors and Computer Vision

Supervisor:

Prof. Ing. Andrea Cereatti

Co-supervisors:

Dr. Ing. Diletta Balta

Dr. Ing. Marco Caruso

Dr. Ing. Daniele Fortunato

Dr. Ing. Rachele Rossanigo

Ing. Paolo Tasca

Candidate:

Mattia Cappello

Academic Year 2025/2026

*Verrà un giorno in cui ciò che ora sfugge
sarà portato alla luce dal tempo
e dalla ricerca di parecchie generazioni.*
— Seneca, *Naturales Quaestiones*, VII.25.4

A me stesso,
per aver trovato la forza nelle avversità.

Alla mia famiglia,
per essere amore e costante nella mia vita.

Ai miei amici, vecchi e nuovi,
per aver dato colore alla mia esistenza.

Abstract

Recently, advances in wearable inertial technologies have enabled real-world digital mobility assessment. However, mobility measures are often aggregated across contexts, potentially masking environmental adaptations and age-related differences which may emerge during complex activities and concurrent tasks requiring continuous perceptual processing, visual exploration, and motor control. To overcome these limitations, head-worn devices integrating inertial sensors and scene-recording cameras can be used to provide simultaneous information on movement and environmental contexts, enabling locomotor characterisation.

This thesis aimed to investigate context- and age-related differences in gait and head-movement characteristics during urban locomotion by developing a pipeline for context identification using egocentric videos acquired through smart glasses.

The study involved 12 healthy participants divided into two groups: 6 younger adults (20–31 years) and 6 older adults (65–80 years). Participants were equipped with a multi-sensor wearable system for movement analysis (INDIP) synchronised with smart glasses for first-person scene recording, during a 45-minute urban walking protocol involving straight walking and pedestrian crossings, with and without face-to-face conversation.

Videos were processed using a semi-automated pipeline based on public Deep-Learning models, such as Places365 and YOLOv11, for crossroad and traffic-light detection. Face-to-face conversations were manually annotated. Stride-by-stride gait parameters were extracted from the INDIP system together with measures of magnitude, smoothness and regularity from the head-mounted inertial sensor.

Generalised linear mixed-effects models were used to evaluate the effects of cohort, locomotor context and conversation. Locomotor context included straight walking, signalised crossing and unsignalised crossing, while conversation distinguished walking with and without face-to-face interaction. Aggregated analyses ignoring locomotor context were also performed to compare context-aware with context-free parameter values.

When data were aggregated across all locomotor contexts, a significant difference between younger and older adults was observed only for vertical head-movement

regularity, quantified by RCME V-HD, with older adults exhibiting higher values. No significant cohort differences emerged for spatiotemporal gait parameters, head-movement magnitude, head-movement smoothness, or the other entropy components.

Context-aware analyses revealed a more articulated pattern of results. All spatiotemporal gait parameters were significantly influenced by locomotor context and by the Context $\times$ Conversation interaction, indicating gait adaptations across different urban locomotor conditions and in relation to the presence of conversation. Froude number and stride speed also showed significant cohort effects, while stride length showed a significant Cohort $\times$ Context interaction. Head-movement parameters showed direction- and metric-specific effects: head-movement magnitude was influenced by context and conversation in the medio-lateral and antero-posterior directions, vertical head-movement regularity consistently distinguished the two cohorts, and head-movement smoothness mainly showed interaction effects involving cohort, context and conversation.

Overall, this work shows that contextual information can reveal context-dependent adaptations and selective age-related differences in gait dynamics and head-motion control which may remain hidden when locomotor events are aggregated.

Acknowledgements

At the end of this work, I would like to express my most sincere gratitude to all the people who have shaped, guided, and supported me throughout the development of this thesis.

First and foremost, I would like to thank Prof. Andrea Cereatti, my supervisor, for believing in me and for giving me this valuable opportunity for both academic and personal growth. Thanks to his trust, I had the chance to challenge myself, gain new experiences, and contribute, in my own small way, to an innovative project with great potential.

I would also like to sincerely thank my co-supervisors, Ing. Diletta Balta, Ing. Marco Caruso, Ing. Daniele Fortunato, Ing. Rachele Rossanigo, and Ing. Paolo Tasca, for their help, availability, and constant support from the very beginning of this journey. Their guidance and encouragement accompanied me throughout the development of this work in the best possible way.

My sincere thanks also go to Dr. Marco Boldregghini for providing part of the instrumentation used in this study and for his valuable support in the recruitment of healthy older adults.

I would also like to thank Ing. Alessandra Audisio for sharing part of this adventure with me and for supporting me along the way.

I would also like to sincerely thank Caterina and Giulia, the interns who, with great commitment and professionalism, helped me carry out this work. Without your valuable contribution, this journey would not have been the same.

Finally, I would like to thank Thea, Milena, Alessio, and Matteo, with whom I shared this thesis journey at the department, for the moments of discussion, support, and companionship that made this experience even more meaningful.

Table of Contents

List of Tables	VII
List of Figures	X
1 Introduction	1
1.1 General Introduction and Motivation	1
1.2 Knowledge Gap	2
1.3 Aim of the Study	3
2 Background	4
2.1 Human Walking and Gait Control	4
2.1.1 Gait Cycle	4
2.1.2 Head and Trunk Movements During Physiological Walking .	9
2.2 Effects of Ageing on Gait and Postural Control	10
2.2.1 Balance Control and Fall Risk in Older Adults	11
2.3 Head Movements and Visual Exploration When Walking	12
2.3.1 Visual Exploration During Locomotion	12
2.3.2 Interaction Between Head Movements and Locomotor Control	13
2.3.3 Ecological Walking and Interactions with the Environment	13
2.3.4 Cognitive Load and Attentional Demands During Locomotion	15
2.4 Wearable Technologies for Ecological Gait Analysis	15
2.4.1 Inertial Measurement Units	15
2.4.2 Smart Glasses and Scene Recording Devices	22
2.4.3 Computer Vision and Egocentric Scene Analysis	23
3 Materials and Methods	24
3.1 Study Participants	26
3.2 Experimental Setup	26
3.2.1 INDIP System	26

3.2.2	Adopted INDIP System Configuration	27
3.2.3	Smart Glasses for Context Recording	29
3.3	Experimental Protocol	31
3.3.1	Out-of-Lab	31
3.4	Data Collection and Processing	33
3.4.1	Data Acquisition	33
3.4.2	Anthropometric Parameters	33
3.4.3	Data Standardisation and Synchronisation	35
3.4.4	Video Annotation and Event Segmentation	39
3.5	Evaluation of the Performance of the Automatic Pipeline	46
3.5.1	Evaluation of Pedestrian Crossing Recognition Performance	46
3.5.2	Evaluation of Traffic Light Recognition Performance	50
3.6	Extraction of Biomechanical Parameters of Gait and Head Movement	54
3.6.1	Preprocessing: Orientation Estimation and Gravity Removal	56
3.6.2	Spatiotemporal Gait Parameters	56
3.6.3	Head-Movement Parameters	57
3.7	Obtained Dataset	60
3.8	Statistical Analysis	62
3.8.1	Generalised Linear Mixed-Effects Models	62
3.8.2	Assessment of Model Adequacy	63
3.8.3	Post-Hoc Analysis	64
4	Results	65
4.1	Performance of Pre-trained Networks without and with Post-processing	65
4.1.1	Performance of Automatic Pedestrian Crossing Recognition	65
4.1.2	Performance of Automatic Recognition of Signalised Crossings	68
4.2	Aggregated Analysis	72
4.3	Context-Aware Analysis	74
4.3.1	Spatiotemporal Gait Parameters	74
4.3.2	Head-Movement Parameters	82
4.4	Post-Hoc Comparisons	95
4.5	Comparison between Aggregated and Context-Aware Analysis	98
5	Discussion	100
5.1	Main findings	100
5.2	Effects of locomotor context on gait parameters	102
5.3	Effects of locomotor context and conversation on head movements	103
5.3.1	Head-movement amplitude	103

5.3.2	Complexity and regularity of head movements	104
5.3.3	Smoothness of head movements	105
5.4	Age-related differences in real-world locomotion	106
5.5	Importance of the context-aware approach	107
5.6	Performance of the Automatic Context-Recognition Pipeline	108
5.7	Methodological Considerations	110
5.8	Study Limitations	110
5.9	Future Perspectives	111
6	Conclusion	113
	Bibliography	115

List of Tables

3.1	Number of participants recruited and included in the final analyses.	26
3.2	Technical specifications of the sensors integrated in the INDIP IMU [50].	27
3.3	Technical specifications of the Tobii Pro Glasses 3 used in this study, adapted from the manufacturer manual.	30
3.4	Anthropometric measurements collected for each participant and used to personalise the INDIP processing algorithms.	34
3.5	Parameters used for the temporal stabilisation of Places365 predictions.	41
3.6	Parameters used for traffic light detection and tracking.	42
3.7	Parameters used for traffic light state classification and temporal stabilisation.	43
3.8	Categories used for event annotation.	45
3.9	Conversation categories annotated during the manual review.	45
3.10	Metrics used to evaluate the performance of pedestrian crossing recognition.	50
3.11	Extracted biomechanical parameters and corresponding physiological meaning.	55
3.12	Number of events and strides included in the statistical analysis, stratified by locomotor context and conversation condition.	61
4.1	Frame-level performance of raw and post-processed Places365 crosswalk detection.	65
4.2	Event-level Places365 crosswalk detection performance using a 20% bidirectional overlap criterion.	66
4.3	Sensitivity analysis of event-level Places365 crosswalk detection across bidirectional overlap thresholds.	67
4.4	Frame-level traffic-light detection performance within crosswalk frames.	68
4.5	Event-level traffic-light classification performance using 20% event threshold.	69
4.6	Sensitivity analysis of event-level traffic-light classification across event thresholds.	69

4.7	Percentage of frames classified with traffic lights for correctly classified traffic-light events (TP) and false negative events (FN).	70
4.8	Aggregated GLMM results.	72
4.9	Froude number [-]. Descriptive values are reported as median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	75
4.10	Stride speed [m/s]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	76
4.11	Stride length [m]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	78
4.12	Stride duration [s]. Descriptive values are reported as median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	79
4.13	Cadence [steps/min]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	81
4.14	aRMS/speed V-HD [s^{-1}]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	83
4.15	aRMS/speed ML-HD [s^{-1}]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	83
4.16	aRMS/speed AP-HD [s^{-1}]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	84
4.17	RCME V-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	87
4.18	RCME ML-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	87
4.19	Head RCME AP-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	88
4.20	LDLJ V-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	91

4.21	LDLJ ML-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	91
4.22	LDLJ AP-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.	92
4.23	Significant model-based post-hoc comparisons after FDR correction.	95
4.24	Comparison between aggregated and context-aware analyses. Only statistically significant effects are reported.	98

List of Figures

2.1	Representation of step and stride during walking. A step corresponds to the interval between the initial contact of one foot and the initial contact of the contralateral foot, whereas a stride corresponds to two consecutive initial contacts of the same foot.	5
2.2	Main phases of the gait cycle, including the subphases of the stance and swing phases. Adapted from [10].	6
2.3	Simplified representation of the human locomotor system divided into two functional components: the <i>Passenger Unit</i> , representing the upper body mass carried when walking, and the <i>Locomotor Unit</i> , representing the lower limbs responsible for movement generation and propulsion during gait.	9
2.4	Experimental setup of a virtual reality street-crossing simulation adapted from [35]. The participant, equipped with an eye-tracking and head-tracking system, stands in front of an immersive environment consisting of three screens reproducing an urban pedestrian-crossing scenario. The three projection screens provide a wide field of view to simulate ecological navigation and pedestrian decision-making processes during crossing tasks.	14
2.5	Illustrative representation of pedestrians crossing a signalised urban intersection during ecological walking. Urban environments require continuous integration of sensory information, attentional processing, environmental exploration, and locomotor adaptation to ensure safe navigation and effective interaction with the surrounding dynamic environment (AI-generated image).	14
2.6	Simplified representation of a mass–spring–damper accelerometric system.	16
2.7	Examples of vibrating tuning-fork gyroscope configurations.	19
2.8	Simplified representation of the Hall effect principle used in magnetometers.	20

2.9	Schematic representation of an infrared proximity sensor based on the <i>time-of-flight</i> principle.	21
2.10	Examples of wearable smart glasses used for egocentric scene recording and visual behaviour analysis.	22
3.1	Overview of the methodological workflow adopted in the study, from the acquisition of INDIP and smart glasses data to synchronisation, context recognition, manual review, biomechanical parameter extraction, event-level aggregation, and statistical analysis.	25
3.2	Structure of the inertial measurement unit (IMU) used in the INDIP system. The external casing, the electronic board containing the inertial sensors and data acquisition circuits, the integrated battery, and the upper cover of the device are visible.	27
3.3	Schematic representation of the INDIP system configuration adopted in the study. The inertial sensor positioned on the head, the module installed at the lumbar level, the inertial units mounted on the footwear, and the sensors used to measure the distance between the feet are highlighted.	29
3.4	Tobii Pro Glasses 3 used in the present study: device (a) and wearing configuration during data acquisition (b).	30
3.5	Standardised out-of-lab route used in the study. All participants followed the same urban itinerary near the Politecnico in Turin, including straight walking segments, pedestrian crossings, and sections dedicated to controlled interactions.	32
3.6	Standard reference system adopted by the INDIP framework for the representation of signals acquired from the IMU units. The convention defines the anatomical antero-posterior (AP), medio-lateral (ML), and vertical (V) axes, ensuring a consistent representation of the data from the different sensors. Adapted from Palmerini et al. [55].	36
3.7	Example of the synchronisation procedure between the INDIP system and the Tobii Pro Glasses 3 based on the comparison of the norm of the angular velocity measured by the gyroscopes mounted on the head. In (a), the signals before temporal delay compensation are shown, whereas in (b), the same signals after temporal alignment are reported.	38
3.8	Scheme of the processing pipeline used by the INDIP framework for gait event identification and the estimation of spatiotemporal parameters. Adapted from Salis et al. [50].	39

3.9	Example of temporal alignment between event annotations obtained from egocentric videos and inertial signals acquired by the INDIP system. The figure shows the correspondence between pedestrian crossing annotations and the medio-lateral gyroscopic signals recorded by the sensors positioned on the feet, used to verify the correct synchronisation between contextual information and biomechanical data.	55
4.1	Trend of the F1-score as a function of the bidirectional overlap threshold between manual and automatically detected events. . . .	67
4.2	Trend of recall as a function of the bidirectional overlap threshold between manual and automatically detected events.	68
4.3	Trend of recall as a function of the threshold used to classify an event as a signalised crossing.	70
4.4	Trend of the F1-score as a function of the threshold used to classify an event as a signalised crossing.	71
4.5	Aggregated distribution of RCME V-HD, the only parameter found to be significantly different between the two cohorts in the aggregated analysis.	73
4.6	Distribution of Froude number values across the different locomotor contexts and conversation conditions, considered separately for each cohort.	75
4.7	Distribution of stride speed values in the different locomotor contexts and conversation conditions, considered separately for each cohort. . .	77
4.8	Distribution of stride length values in the different locomotor contexts and conversation conditions, considered separately for each cohort. . .	78
4.9	Distribution of stride duration values in the different locomotor contexts and conversation conditions, considered separately for each cohort.	80
4.10	Distribution of cadence values in the different locomotor contexts and conversation conditions, considered separately for each cohort. . .	81
4.11	Distribution of aRMS/speed values in the different locomotor contexts and conversation conditions, considered separately for each cohort: (a) vertical component, (b) medio-lateral component, and (c) anterior-posterior component.	85
4.12	Distribution of RCME values across the different locomotor contexts and conversation conditions, considered separately for each cohort: (a) vertical component, (b) medio-lateral component, and (c) anterior-posterior component.	89

4.13 Distribution of LDLJ values in the different locomotor contexts and conversation conditions, considered separately for each cohort: (a) vertical component, (b) medio-lateral component, and (c) anterior-posterior component.	93
---	----

Chapter 1

Introduction

1.1 General Introduction and Motivation

In recent years, the study of human behaviour in real-world environments has become increasingly important in several research fields, including telerehabilitation and telemonitoring. In these contexts, gait analysis represents a fundamental tool to understand how individuals adapt their motor strategies while interacting with the surrounding environment. Gait, also recognised as the sixth vital sign [1], provides valuable information about an individual's health status. Changes associated with physiological ageing or the progression of neurological disorders such as Parkinson's disease and multiple sclerosis can also be observed.

However, walking does not depend solely on biomechanical factors. The way a person moves is strongly influenced by the environment and by the sensory stimuli received during movement. For this reason, gait can be considered the result of the continuous interaction between motor and cognitive components, especially in situations which require the management of multiple tasks simultaneously, a condition commonly referred to as *dual-tasking* or, more generally, *multi-tasking*.

In urban environments, pedestrian crossings require individuals to constantly adjust their walking behaviour according to traffic-light signals, approaching vehicles, surrounding noise, and the presence of other pedestrians. These external factors can affect both gait and attention, particularly in older adults, who may experience greater difficulty managing simultaneous cognitive and motor demands.

For this reason, analysing gait parameters alone without considering the environmental context may lead to incorrect interpretations. Indeed, some gait variations may simply represent adaptations to the surrounding conditions rather than being indicators of clinical deterioration. For example, in crowded environments, a reduction of walking speed accompanied by shorter steps compared with those recorded during standard unobstructed free-walking conditions can be commonly observed

[2].

1.2 Knowledge Gap

Numerous studies on human gait have been reported in literature. Most of them are based on laboratory experiments, where variables and the level of complexity can be controlled with high precision. Although this represents a methodological advantage, it also makes it more difficult to observe different types of behaviour which are truly representative of everyday life.

Pedestrian crossing is one among the most extensively studied contexts. This scenario is characterised by a large number of stimuli originating from the surrounding environment. Factors such as traffic density, the number of people present, traffic light colour, and environmental noise must be continuously processed in order to adopt the most appropriate motor strategy for the situation. For example, if the traffic lights are green but the siren of an emergency vehicle can be heard in the distance, individuals are likely to pay greater attention to the surrounding environment to avoid obstructing its path. This increased attention may consequently be reflected in a reduction in walking speed.

To expose individuals to conditions which resemble real-world situations, several studies have introduced simulators and virtual reality environments in which the level of complexity can be modified [3, 4]. These approaches reproduce controlled scenarios characterised by different levels of complexity. However, some key aspects are difficult to replicate, including the unpredictability of events, risk perception, and the level of sensory involvement which is typical of real-world environments. Furthermore, in these contexts, individuals are often required to take decisions while considering their potential consequences. This is difficult to reproduce in simulated environments.

A similar situation applies to *dual-task* paradigms, which are widely used to investigate the interaction between cognitive and motor components. In many cases, the secondary task consists of abstract activities, such as memory exercises or mental calculations [5], which hardly ever reflect the cognitive demands of everyday situations.

In recent years, some studies have attempted to shift the focus towards more ecological contexts, highlighting the role of sensory information in pedestrian decision-making processes, for example in estimating the time available to cross a road [6].

Although some studies have investigated visual behaviour and decision-making processes during pedestrian crossing [7, 8], the relationship between locomotion, visual attention, and environmental dynamics remains only partially understood. In particular, studies capable of simultaneously acquiring motor data, visual

information, and environmental context information during complex ecological scenarios are still limited.

1.3 Aim of the Study

In light of the limitations highlighted above, this work aims to investigate gait and head-movement characteristics during real-world urban locomotion, with particular attention to the role of environmental context and concurrent interaction.

The focus of the present study is on situations in which pedestrians are required to adapt their behaviour to changing environmental conditions, including traffic, traffic-light signals, and interactions with other road users. Pedestrian crossing was selected as the main scenario of interest because it requires individuals to process information from the surrounding environment and adjust their movements accordingly.

To investigate these types of behaviour, wearable inertial sensors and scene-recording smart glasses were used to simultaneously collect movement and visual information during real-world walking. This approach enabled the analysis of pedestrian behaviour in ecological conditions and provided quantitative measures of the strategies adopted during everyday mobility tasks.

Chapter 2

Background

2.1 Human Walking and Gait Control

Human mobility is closely related to locomotion, namely the ability to voluntarily move from one place to another. It represents an essential function in everyday life, as it allows individuals to interact with the surrounding environment and carry out daily activities. In humans, walking is the most common form of natural terrestrial locomotion. Its role is to ensure body progression while maintaining stability and efficiency, allowing movement across different surfaces with the lowest possible energy expenditure [9].

Although walking is generally considered an automatic action, it actually involves numerous processes which must operate in a coordinated manner. During locomotion, the nervous system continuously integrates motor, sensory, and cognitive information in order to maintain balance, control movement, and adapt motor behaviour to the characteristics of the surrounding environment. For this reason, walking cannot be considered a simple repetitive movement, but rather a dynamic and complex process regulated by multiple neuromuscular and biomechanical mechanisms.

2.1.1 Gait Cycle

Human gait can be described as a cyclic and quasi-periodic phenomenon, in which motor sequences are repeated over time in a similar manner [9]. This characteristic is particularly evident when walking at a constant speed, a condition in which movement tends to maintain a regular and stable pattern. Conversely, during gait initiation, stopping phases, or speed changes, the locomotor pattern becomes less regular.

Under physiological conditions, locomotor control during steady-state walking mainly occurs at the spinal level through neuronal mechanisms responsible for the

rhythmic generation of movement, known as central pattern generators (CPGs). In the presence of neurological or musculoskeletal disorders, this regularity may decrease, resulting in a less stable and efficient gait [9].

Gait analysis is generally performed in laboratory settings under standardised conditions, with straight trajectories, flat surfaces, and constant gait speed, so that the assumption of quasi-periodicity can be considered valid. In this way, it is possible to observe a limited number of gait cycles and statistically describe their main characteristics, such as mean values and variability. This approach allows the identification of motor alterations and supports clinical evaluation, surgical planning, orthotic prescription, and rehabilitation treatment.

The basic unit of gait is the *step*, defined as the interval between the contact of one foot with the ground and the subsequent contact of the contralateral foot. Two consecutive steps, one right and one left, form a *stride*, or gait cycle. A stride corresponds to the interval between two consecutive initial contacts of the same foot, as shown in Figure 2.1.

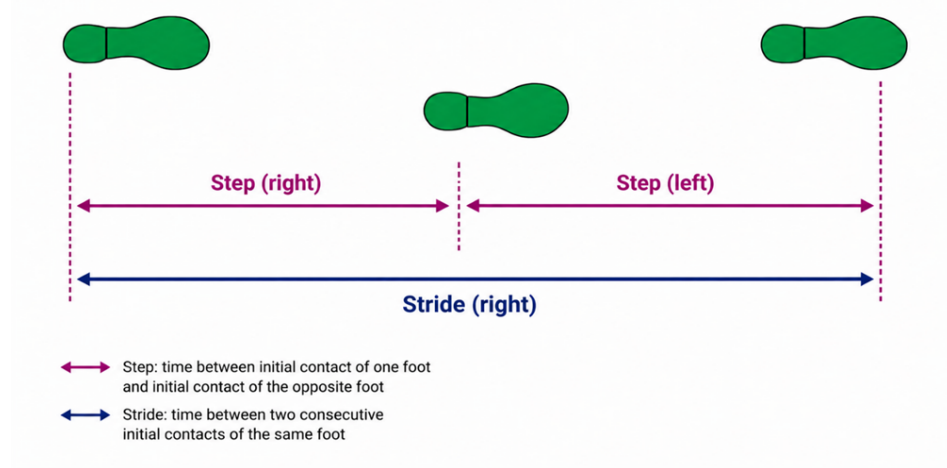


Figure 2.1: Representation of step and stride during walking. A step corresponds to the interval between the initial contact of one foot and the initial contact of the contralateral foot, whereas a stride corresponds to two consecutive initial contacts of the same foot.

Phases of the Gait Cycle

The gait cycle is traditionally divided into two main phases: the *stance phase* and the *swing phase*.

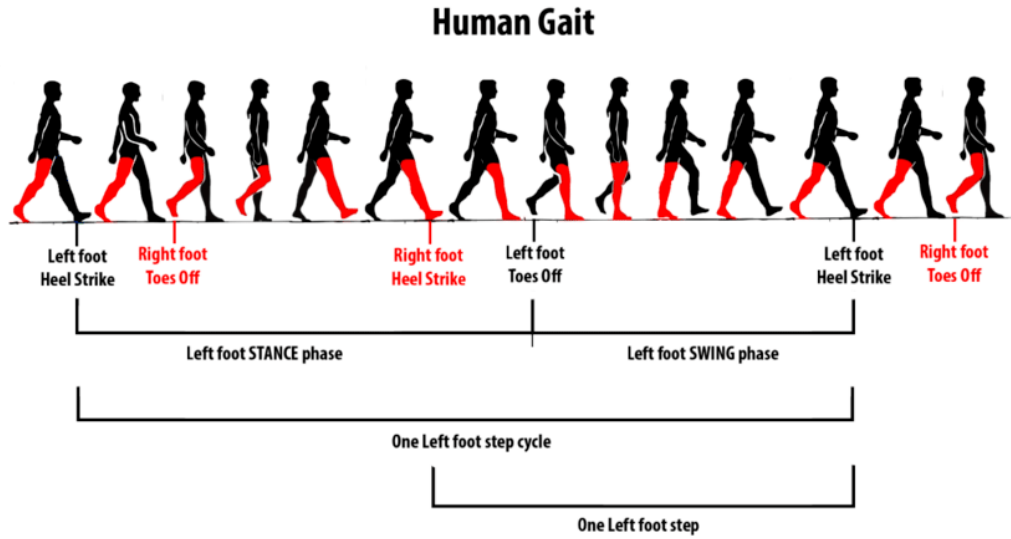


Figure 2.2: Main phases of the gait cycle, including the subphases of the stance and swing phases. Adapted from [10].

Stance Phase

The stance phase represents approximately 60% of the gait cycle and begins when the foot makes contact with the ground. During this phase, the lower limb provides body support, contributes to balance control, absorbs impact forces, and allows forward progression. [11]. According to literature, the stance phase can be divided into five subphases [12].

- **Initial Contact (Heel Strike) [0–2%]**

The *initial contact* represents the moment when the foot first comes into contact with the ground, typically through the heel. During this subphase, the limb is positioned for weight acceptance and for the initial rolling motion of the foot [9].

- **Loading Response (Foot Flat) [2–10%]**

During the loading response, body weight is progressively transferred onto the supporting limb. The knee flexes slightly to absorb the impact with the ground and reduce joint stress. This phase also includes the first period of double support, when both feet are simultaneously in contact with the ground [9, 11, 12].

- **Mid Stance [10–30%]**

During the mid-stance subphase, the body progresses over the supporting limb through the inverted pendulum mechanism, which allows energy-efficient progression of the centre of mass. In this interval, the foot represents the only point of contact with the ground, making lower-limb and trunk stability essential [9, 11, 12, 13].

- **Terminal Stance (Heel Off) [30–50%]**

During terminal stance, the heel begins to rise from the ground while the body continues its forward progression. Propulsive force is generated through the activation of the ankle plantarflexor muscles, particularly the triceps surae. Body weight is progressively transferred towards the forefoot in preparation for foot-off [13].

- **Pre-Swing (Toe Off) [50–60%]**

This phase represents the transition between stance and swing. Body weight is transferred to the contralateral limb while the knee of the reference limb begins to flex, preparing the foot to leave the ground initiating limb advancement.

Swing Phase

The swing phase accounts for approximately 40% of the gait cycle and corresponds to the period when the foot is not in contact with the ground. Its primary purpose is to advance the limb forward and prepare for the subsequent initial contact [11]. The swing phase can be divided into three main subphases.

- **Initial Swing [60–73%]**

This phase begins when the foot leaves the ground. Its main objective is to achieve adequate foot clearance and initiate advancement of the lower limb [9].

- **Mid Swing [73–87%]**

The limb continues moving forward until the tibia reaches an approximately vertical position [12]. During this phase, maximum knee flexion is generally observed, allowing the foot to avoid contact with the ground. Part of the movement takes advantage of limb inertia, making walking more energy efficient [9].

- **Terminal Swing [87–100%]**

The limb progressively decelerates and prepares for the subsequent ground contact. The lower-limb muscles act to control knee extension and properly position the foot in preparation for the next gait cycle [9, 13].

Spatio-Temporal Gait Parameters

Quantitative gait analysis can be performed using local parameters, related to the behaviour of specific joints or muscle groups, and global parameters, used to describe the locomotor pattern as a whole. Among the latter, spatiotemporal parameters are one of the most widely used tools for gait assessment, as they allow the description of aspects related to progression, stability, and movement efficiency.

The main spatiotemporal parameters include stride length, stride duration, and gait speed.

Stride Length

Stride length corresponds to the distance between two consecutive initial contacts of the same foot during the gait cycle. This parameter depends on several factors, including lower-limb length, joint mobility, muscle strength, and the ability to maintain balance during progression [14]. Reductions in stride length may be associated with compensatory strategies adopted to increase stability or with neurological and musculoskeletal impairments [15]. In clinical settings, this parameter is frequently used to assess an individual's level of mobility and the effectiveness of rehabilitation interventions.

Stride Duration

Stride duration represents the time interval between two consecutive initial contacts of the same foot. It is closely related to stride frequency, of which it is the mathematical inverse ($\text{Stride Duration} = 1/\text{Stride Frequency}$).

A reduction in stride duration generally corresponds to an increase in stride frequency and, consequently, to a higher gait speed [14]. The analysis of temporal gait parameters also provides information about motor coordination and gait symmetry. Alterations in stride duration may reflect compensatory strategies or deficits in neuromuscular control [16].

Gait Speed

Gait speed, generally expressed in metres per second (m/s), describes the speed at which an individual moves through space. It represents one of the main indicators of locomotor performance [1] and it is influenced by both stride length and stride frequency. Under physiological conditions, an adequate gait speed is associated with energy-efficient movement [17]. Conversely, a significant reduction in gait speed may indicate pain, muscle weakness, neurological deficits, or the need to increase stability during walking [14]. Particularly high values may instead reflect a less controlled gait pattern or different movement strategies.

2.1.2 Head and Trunk Movements During Physiological Walking

When walking, the body can be considered as consisting of two main functional units with different roles: the *locomotor unit* and the *passenger unit* [14].

The locomotor unit includes the pelvis and lower limbs and it is primarily responsible for body support and forward progression. In addition to generating movement, it also contributes to shock absorption during foot-ground contact, maintenance of vertical stability, and reduction of energy expenditure during locomotion [14].

The passenger unit includes the trunk, head, and upper limbs. Its role is to limit excessive oscillations and movements of the upper body during progression. In particular, the oscillatory movement of the arms helps compensate for part of the rotations generated by the lower limbs, contributing to a more stable and efficient gait [14].

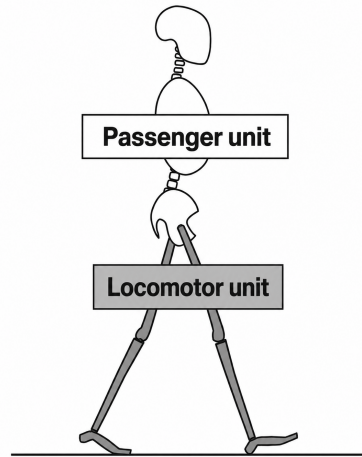


Figure 2.3: Simplified representation of the human locomotor system divided into two functional components: the *Passenger Unit*, representing the upper body mass carried when walking, and the *Locomotor Unit*, representing the lower limbs responsible for movement generation and propulsion during gait.

The head must also remain sufficiently stable to ensure proper perception of the surrounding environment. Head stabilisation is achieved through continuous coordination between head and trunk movements, together with the integration of visual, vestibular, and proprioceptive information [18, 19, 20]. This allows gaze to remain stable even while the body is moving.

During physiological walking, head movements compensate for part of the oscillations generated by the rest of the body in both the sagittal and horizontal

planes [18, 20]. The trunk also contributes to this mechanism by progressively attenuating the accelerations which propagate from the lower limbs towards the head [20, 21]. As a result, head oscillations are generally smoother and more controlled than those observed at the trunk level [21].

Body oscillations also vary according to the direction considered. In the vertical and anteroposterior directions, two main perturbations are observed during each stride, corresponding to the two steps. In the mediolateral plane, however, movement follows a different pattern characterised by a single right-to-left oscillation during the gait cycle [22].

Overall, coordination between the head and trunk plays a fundamental role in maintaining balance, stabilising sensory information, and ensuring efficient walking, particularly in complex environments or situations that require continuous motor adaptations [21].

2.2 Effects of Ageing on Gait and Postural Control

Walking is an essential component of human mobility and it is closely linked to functional independence and quality of life. The ability to move safely and effectively allows individuals to perform daily activities, maintain social relationships, and interact with the surrounding environment. During ageing, changes in motor and functional abilities may compromise these capabilities, with significant consequences for individual autonomy.

In clinical and scientific settings, individuals aged 65 years and older are generally classified as older adults, although the ageing process shows considerable inter-individual variability. From this age onwards, several physiological changes which may affect gait, postural control, and functional mobility become progressively more evident.

Human locomotion is a complex process that requires continuous integration between musculoskeletal, sensory, and cognitive systems. Age-related alterations affecting one or more of these systems may reduce locomotor efficiency and modify the strategies adopted during walking [14]. Among the main factors responsible for these changes are reduced muscle strength, slower motor responses, and declining sensory abilities. In particular, lower-limb weakness may reduce propulsion during gait and limit the ability to maintain high walking speeds [23].

Among the spatiotemporal parameters most commonly used to describe functional mobility, gait speed is considered one of the most relevant indicators. Several studies have reported a progressive decline in gait speed with advancing age, particularly after the age of 70. Reduced gait speed is also associated with lower functional capacity, an increased risk of hospitalisation, and higher mortality among

older adults [23, 24].

Ageing is also associated with changes of other spatiotemporal gait parameters. In particular, older adults frequently exhibit a reduction of stride length and an increase in double-support duration compared with younger individuals [25]. These changes are generally interpreted as locomotor adaptations aimed at increasing safety during walking.

Gait alterations tend to become even more evident in complex or cognitively demanding situations, in which continuous adaptation to environmental conditions is required. In these contexts, older adults often show greater gait variability and a reduced capacity for locomotor adaptation compared with younger individuals [26, 27].

2.2.1 Balance Control and Fall Risk in Older Adults

Maintaining balance when walking relies heavily on the effectiveness of the visual, vestibular, and proprioceptive systems, which allow accurate perception of body position and movement in space. With advancing age, subclinical or pathological decline of these sensory systems impair postural control, reducing dynamic stability during locomotion [28, 29].

To compensate for sensory uncertainty and the resulting instability, the central nervous system adopts more conservative walking strategies. This results in an exacerbation of the previously described spatiotemporal adaptations—such as shorter steps and prolonged double-support duration—and in a progressive and increasing reliance on visual information (visual anchoring) to compensate for proprioceptive and vestibular deficits [7, 25, 30].

These deficits and the associated compensatory strategies become critical in ecological or ecologically valid contexts, where walking does not occur in isolation but requires continuous adaptation to external stimuli and the simultaneous management of cognitive tasks (dual-tasking). Everyday situations such as obstacle negotiation, walking in crowded environments, or street crossing drastically increase attentional demands, may cognitively fatigue older adults, and lead to greater gait variability accompanied by a reduced capacity for locomotor adaptation compared with younger individuals [26, 27, 31].

For these reasons, fall risk cannot be attributed solely to a motor component but rather emerges from the multifactorial interaction between sensory efficiency, residual attentional capacity, and environmental complexity. Understanding how these systems cooperate and where they fail is essential for analysing real-world locomotor behaviour and developing effective prevention strategies.

2.3 Head Movements and Visual Exploration When Walking

Human locomotion in real-world environments requires continuous interaction between motor control, sensory processing, and perception of the surrounding context. While walking individuals must constantly adapt their movements according to the environment. They must simultaneously maintain balance, stability, and efficient progression.

In addition to lower-limb activity, head movements also play a fundamental role during locomotion. In particular, head movements contribute to both postural control and environmental exploration, allowing individuals to orient their attention, acquire visual information, and adjust their behaviour in response to external stimuli [32, 33].

These mechanisms become particularly important when walking in dynamic and visually complex environments, where individuals must continuously integrate sensory information, plan movement, and adapt to changes in the surrounding context [31, 34]. Age-related changes affecting postural control, sensory integration, and attentional processes can profoundly alter the way older adults interact with the surrounding environment during locomotion. For this reason, the analysis of head movements when walking can provide useful information about the functional adaptations of mobility and the strategies adopted for interaction with the environment.

2.3.1 Visual Exploration During Locomotion

While walking, individuals continuously acquire and process information from the surrounding environment in order to identify obstacles, estimate distances, plan trajectories, and adapt movement to external conditions [32, 35]. This process does not exclusively involve eye movements but it also includes coordinated movements of the head and body, which allow individuals to expand their field of view and direct attention towards relevant elements of the environment [33, 32].

Visual exploration plays a particularly important role in ecological contexts characterised by high environmental complexity, such as crowded urban areas, pedestrian crossings, or situations which require frequent locomotor adaptations. In these scenarios, the ability to rapidly gather visual information and integrate it with movement control is essential for ensuring safe and effective navigation.

Several studies have shown that ageing can modify visual exploration strategies and interaction with the environment during locomotion. Older adults often exhibit greater attentional demands and a reduced capacity for adaptation in complex environments, particularly under multitasking conditions or in the presence of high cognitive loads [26, 36].

2.3.2 Interaction Between Head Movements and Locomotor Control

Head movements are closely linked to gait dynamics. The coordination between the head, trunk, and lower limbs responds to a dual requirement: stabilising the visual field and ensuring dynamic balance control through continuous anticipatory and compensatory postural adjustments [37, 35].

During complex motor tasks, such as changes in direction or obstacle negotiation, head rotations undergo significant kinematic variations which reflect both trajectory planning and pedestrian decision-making processes [32, 38]. In older adults, however, the ability to coordinate these visuomotor stimuli with the demands of walking is substantially reduced when the environment requires continuous adaptation to irregular surfaces or dynamic contexts.

Under *dual-task* conditions, the introduction of a secondary cognitive load exacerbates further this vulnerability, altering the modulation of spatiotemporal gait parameters and increasing the potential risk of tripping or falling [39].

2.3.3 Ecological Walking and Interactions with the Environment

Walking in real-world ecological contexts differs substantially from standardised laboratory locomotion, requiring continuous adaptation to dynamic, heterogeneous, and inherently unpredictable scenarios [31, 34]. Among these contexts, street crossing represents one of the most complex situations: pedestrians must simultaneously monitor traffic flow, estimate the speed of approaching vehicles, comply with the timing constraints of traffic light systems, and maintain walking stability [40, 41].

In urban intersections, due to the high number of stimuli, environmental exploration through head movements becomes a prerequisite for safety. To analyse this behaviour while preserving experimental repeatability, research has made extensive use of immersive virtual reality (VR) combined with eye- and head-tracking systems (Figure 2.4) [35].

However, simulated environments present intrinsic limitations in reproducing the unpredictability of outdoor settings (e.g., changes in lighting conditions, spontaneous social interactions, and unexpected physical obstacles) [34]. For this reason, kinematic analysis conducted directly in real outdoor environments (Figure 2.5) is essential to understand the actual interaction between head kinematics and gait parameters in both younger and older populations.

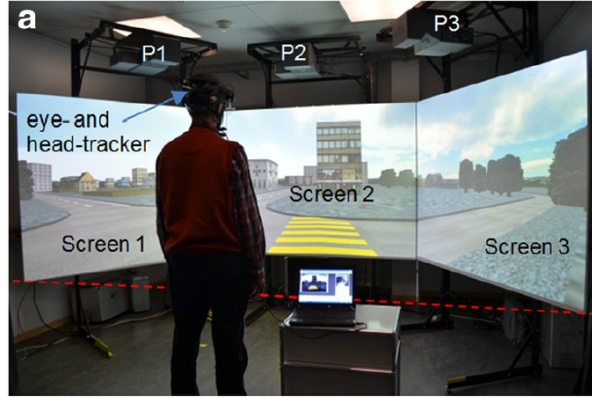


Figure 2.4: Experimental setup of a virtual reality street-crossing simulation adapted from [35]. The participant, equipped with an eye-tracking and head-tracking system, stands in front of an immersive environment consisting of three screens reproducing an urban pedestrian-crossing scenario. The three projection screens provide a wide field of view to simulate ecological navigation and pedestrian decision-making processes during crossing tasks.



Figure 2.5: Illustrative representation of pedestrians crossing a signalised urban intersection during ecological walking. Urban environments require continuous integration of sensory information, attentional processing, environmental exploration, and locomotor adaptation to ensure safe navigation and effective interaction with the surrounding dynamic environment (AI-generated image).

2.3.4 Cognitive Load and Attentional Demands During Locomotion

In everyday life, walking rarely occurs as an isolated motor task. Navigation requires a continuous reallocation of attentional resources between gait control and the execution of concurrent cognitive tasks (*dual-task* or *multi-task*).

Ecological conditions such as social interaction (e.g., conversing with a companion) or smartphone use (e.g., reading or texting messages while walking) drastically increase cognitive load. This overload is immediately reflected in gait biomechanics, resulting in a marked reduction in walking speed, deviations from a straight trajectory, and alterations in gait rhythm [39, 27].

At the same time, attentional demands during locomotion may also influence head movements and environmental exploration strategies. Situations involving social interactions, phone conversations, or complex urban navigation may require continuous reallocation of attentional resources between locomotion and the processing of external stimuli.

2.4 Wearable Technologies for Ecological Gait Analysis

The development of *wearable* technological devices has considerably expanded the possibility of studying human locomotion in ecological and real-world conditions. Unlike traditional motion capture systems used in laboratory settings, wearable devices, thanks to their portability and autonomy, make it possible to investigate locomotor behaviour in complex real-world scenarios characterised by dynamic interaction, terrain variability, and continuous demands for motor adaptation.

In recent years, research has seen the progressive introduction of a wide range of *wearable* sensors, including inertial measurement units, *smart glasses*, and multimodal sensing platforms. The integrated use of these devices enables the simultaneous acquisition of kinematic, behavioural, and contextual information when walking, providing a comprehensive view of movement outside laboratory environments.

2.4.1 Inertial Measurement Units

Inertial measurement units (IMUs) are among the most widely used *wearable* sensors for human movement analysis. They generally integrate accelerometers, gyroscopes, and magnetometers to estimate body accelerations, angular velocities, and segment orientations during movement [42, 43].

Compared with traditional optical motion capture systems, IMU-based systems

offer several advantages for ecological locomotion analysis. Their portability, low cost, and ease of use allow gait assessments to be performed outside laboratory settings and during everyday activities [44]. Furthermore, inertial sensors can be positioned on different body segments, enabling the simultaneous analysis of lower-limb kinematics, trunk motion, and head movements when walking.

IMUs have been widely used for the assessment of spatiotemporal gait parameters, postural control, stability, and locomotor adaptations in both healthy and clinical populations [42]. Older adults are among the populations most frequently investigated using inertial sensors. The possibility of analysing gait in real-world contexts allows a better understanding of how the normal ageing process influences mobility, spatiotemporal gait parameters, and the strategies adopted to adapt to situations characterised by high environmental complexity.

Modern wearable IMUs generally consist of accelerometers, gyroscopes, and magnetometers based on micro-electromechanical system (MEMS) technology, which enables the miniaturisation of sensing units, resulting in reduced dimensions, low power consumption, and relatively non-invasive wearable configurations.

Accelerometers

Accelerometers are sensors designed to measure the linear accelerations generated during body movement. From a mechanical perspective, accelerometers can be approximated as second-order mass–spring–damper systems consisting of an inertial mass suspended by elastic and dissipative elements.

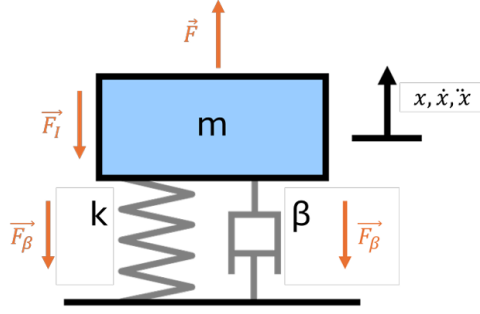


Figure 2.6: Simplified representation of a mass–spring–damper accelerometric system.

When the sensor is subjected to acceleration, the inertial mass tends to oppose the change in motion due to inertia, generating a relative displacement compared to the sensor frame. This displacement is subsequently converted into an electrical signal proportional to the applied acceleration.

According to Newtonian mechanics, the dynamic equilibrium of the inertial mass can be expressed as:

$$F - \beta\dot{x} - kx = m\ddot{x} \quad (2.1)$$

where m represents the inertial mass, β the damping coefficient, k the elastic constant of the system, and x the displacement of the mass.

Since the inertial force applied is proportional to the acceleration experienced by the sensor, Equation 2.1 can be rewritten as:

$$m\ddot{x} + \beta\dot{x} + kx = ma_c \quad (2.2)$$

where a_c represents the external acceleration applied to the accelerometer.

Accelerometers measure the so-called *proper acceleration*, namely the acceleration experienced relative to free fall. When the measurement axis is aligned with the gravitational direction, the acquired signal therefore also includes the contribution of gravity:

$$\vec{a}_p = \vec{a}_s - \vec{g} \quad (2.3)$$

where $\vec{a}_p$ represents the measured acceleration, $\vec{a}_s$ the sensor acceleration, and $\vec{g}$ the gravitational acceleration vector.

The dynamic behaviour of the accelerometer can also be described in the Laplace domain through the transfer function relating the displacement of the inertial mass to the input acceleration:

$$H(s) = \frac{x(s)}{a(s)} = \frac{1}{s^2 + \frac{\beta}{m}s + \frac{k}{m}} \quad (2.4)$$

The transfer function describes the sensitivity of the system to different frequencies. In wearable applications, accelerometers are generally operated well below their resonance frequency in order to ensure stable and reliable measurements during locomotion.

The dynamic response of the sensor depends particularly on the resonance frequency and damping properties. Rearranging Equation 2.4, the transfer function can also be expressed as:

$$H(s) = \frac{1}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \quad (2.5)$$

where ω_0 represents the angular resonance frequency and Q the quality factor of the system.

The resonance frequency can be defined as:

$$\omega_0 = \sqrt{\frac{k}{m}} \quad (2.6)$$

while the quality factor is expressed as:

$$Q = \frac{m\omega_0}{\beta} \quad (2.7)$$

The resonance frequency significantly influences the dynamic behaviour of the accelerometer. Higher resonance frequencies generally allow larger bandwidths to be achieved, although they may reduce measurement sensitivity. For this reason, wearable accelerometers are designed as a compromise between sensitivity, stability, and bandwidth in order to ensure reliable measurements during human movement analysis.

Accelerometers can employ different sensing principles, including capacitive, piezoelectric, piezoresistive, and optical configurations. Capacitive accelerometers are currently the most widely used in wearable technologies due to their small size, low power consumption, and good sensitivity characteristics.

Gyroscopes

Gyroscopes are sensors designed to measure angular velocity during movement. In gait analysis, these sensors are commonly used to investigate body segment rotations, trunk dynamics, postural adaptations, and head movement behaviour during locomotion. Angular velocity measurements are generally expressed in degrees per second ($^\circ/\text{s}$) or radians per second (rad/s).

Most gyroscopes operate according to the Coriolis principle and are commonly referred to as *Coriolis vibrating gyroscopes* or *Vibrating fork gyroscopes*. Within the sensing structure, vibrating masses oscillate at a predefined frequency. When the sensor is subjected to rotation, the Coriolis force generates an additional displacement orthogonal to the vibration direction. This displacement is subsequently converted into an electrical signal proportional to the applied angular velocity.

The Coriolis force acting on the vibrating mass can be expressed as:

$$\vec{F}_c = 2m(\vec{v} \times \vec{\omega}) \quad (2.8)$$

where m represents the inertial mass, $\vec{v}$ the velocity vector of the vibrating mass, and $\vec{\omega}$ the angular velocity vector.

Vibrating tuning-fork gyroscopes generally include pairs of masses oscillating in opposite directions in order to increase sensitivity. When the sensor is subjected to rotational motion, the vibrating masses tend to oscillate along an orthogonal direction due to the Coriolis effect, generating measurable torsional displacements within the sensing structure.

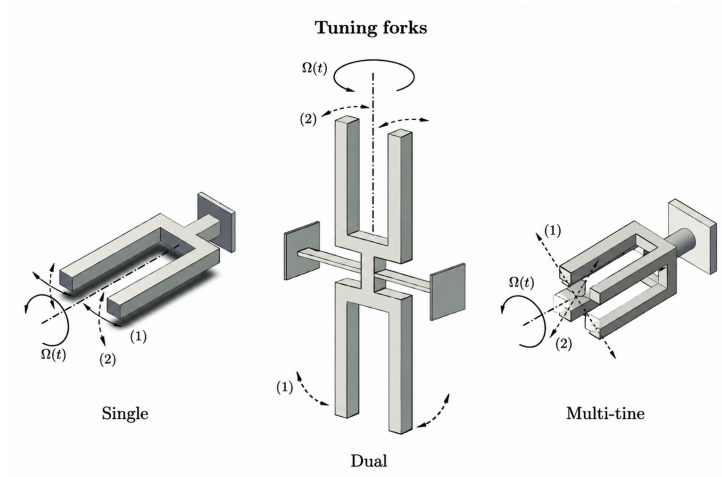


Figure 2.7: Examples of vibrating tuning-fork gyroscope configurations.

From a dynamic perspective, gyroscopes can be approximated as second-order vibrating systems. The mechanical behaviour of the sensing structure can be described as:

$$m\ddot{x} + \beta\dot{x} + kx = F_c \quad (2.9)$$

where m represents the inertial mass, β the damping coefficient, k the elastic constant of the vibrating structure, and F_c the Coriolis force generated by rotational motion.

The dynamic response of the gyroscope mainly depends on the resonance frequency and damping properties. Gyroscopes are generally designed to operate close to resonance in order to increase sensitivity to rotations while maintaining good measurement stability.

Magnetometers

Magnetometers are sensors designed to measure magnetic fields. In wearable systems, these sensors are commonly used to estimate body orientation and movement direction through the detection of the Earth's magnetic field.

Most magnetometers generally operate according to electromagnetic principles such as the Hall effect. When electrically charged particles move within a conductive material in the presence of a magnetic field, they experience a force which is orthogonal to both the magnetic field direction and the particle velocity. This interaction generates a potential difference proportional to the intensity of the magnetic field.

The Lorentz force acting on a charged particle can be expressed as:

$$\vec{F}_L = q(\vec{v} \times \vec{B}) \quad (2.10)$$

where q represents the electric charge, $\vec{v}$ the particle velocity vector, and $\vec{B}$ the magnetic flux density vector.

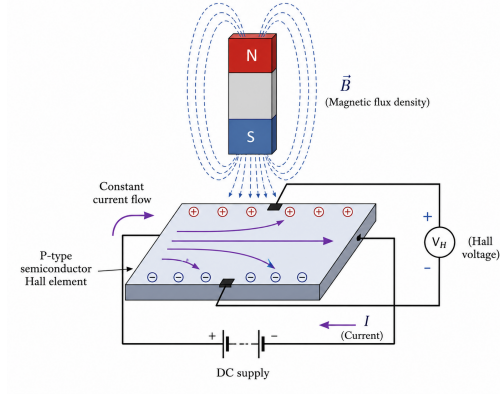


Figure 2.8: Simplified representation of the Hall effect principle used in magnetometers.

Under equilibrium conditions, the Lorentz force is balanced by the electric force generated within the sensing structure:

$$\vec{F}_{TOT} = q(\vec{v} \times \vec{B}) + q\vec{E}_{Hall} \quad (2.11)$$

When the system reaches equilibrium:

$$(\vec{v} \times \vec{B}) = -\vec{E}_{Hall} \quad (2.12)$$

where $\vec{E}_{Hall}$ represents the electric field generated by the Hall effect.

The potential difference produced within the sensing element is proportional to the magnetic field intensity and it can therefore be used to estimate sensor orientation with respect to the Earth's magnetic field.

In wearable inertial systems, magnetometers are commonly integrated with accelerometers and gyroscopes through sensor fusion algorithms in order to improve orientation estimation and compensate for gyroscope drift during prolonged acquisitions.

Pressure Insoles

Pressure insoles are wearable systems designed to measure plantar pressure distribution during standing and walking. These devices generally consist of arrays of

force-sensitive sensors integrated into flexible insoles which can be directly placed inside the subject's footwear.

The sensing elements continuously detect the force exerted by the foot on the ground during locomotion, allowing the acquisition of temporal and spatial information related to foot-ground interaction throughout the gait cycle.

Pressure insole systems can provide information regarding plantar pressure distribution, centre of pressure progression, stance and swing phases, load transfer, and gait event detection. In particular, these systems are commonly used to identify gait-specific events such as heel strike and toe-off when walking.

Depending on the technology employed, plantar pressure measurement systems may be based on resistive, capacitive, or piezoelectric mechanisms. Among these, force-sensitive resistive sensors are among the most widely used in wearable applications due to their flexibility, compact size, and ease of integration into portable systems.

Proximity Sensors

Proximity sensors are devices designed to estimate the distance between the sensor and surrounding objects without requiring direct physical contact. In gait analysis, these sensors can be used for locomotor event detection and for the measurement of spatial gait parameters, such as the distance between the lower limbs during walking (step width) [45], providing useful information about the interaction between the individual and the surrounding environment.

Among the technological devices the most commonly used in wearable systems are infrared *time-of-flight* (ToF) sensors. These devices estimate distance by analysing the characteristics of the infrared signal reflected by the target object. In particular, the measurement can be obtained by evaluating either the signal propagation time or the phase shift between the emitted and received waves.

In modern ToF sensors, the emitter and receiver are generally integrated within the same component, allowing a significant reduction in device size and facilitating their integration into wearable systems for human movement monitoring.

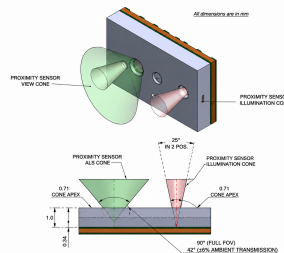


Figure 2.9: Schematic representation of an infrared proximity sensor based on the *time-of-flight* principle.

2.4.2 Smart Glasses and Scene Recording Devices

Wearable devices for scene recording are increasingly being used in ecological locomotion analysis in order to acquire visual information regarding both human behaviour and the surrounding environmental conditions when walking. To maintain good ecological validity, wearable acquisition systems should interfere as little as possible with spontaneous locomotion and daily activities while ensuring adequate portability and comfort.

Smart Glasses

In recent years, smart glasses have gained increasing popularity as wearable platforms for data acquisition during everyday activities. These devices generally integrate miniaturised cameras, microphones, wireless communication systems, inertial sensors, and, in some cases, eye-tracking technologies within lightweight head-mounted structures.

Smart glasses enable the recording of visual scenes directly from the user's perspective (egocentric vision), providing a representation of the environment that is closely related to the perceptual experience during locomotion. This characteristic is particularly relevant for the study of visual exploration, attentional behaviour, obstacle avoidance, and pedestrian-environment interaction in ecological walking contexts.

Thanks to the possibility of combining video information with data acquired from integrated sensors, these devices can support the simultaneous analysis of locomotor behaviour and surrounding environmental conditions. However, they may present some limitations related to battery life, motion artefacts, variability in lighting conditions, and acquisition quality during prolonged outdoor recordings.



(a) Ray-Ban Meta smart glasses



(b) Tobii Pro Glasses 3 wearable eye-tracking system

Figure 2.10: Examples of wearable smart glasses used for egocentric scene recording and visual behaviour analysis.

2.4.3 Computer Vision and Egocentric Scene Analysis

Recent developments in the field of computer vision have significantly improved the automatic analysis of acquired visual information. In recent years, deep-learning approaches have been increasingly used for scene recognition, object detection, activity analysis, and contextual understanding from recordings obtained through wearable cameras [46, 47].

Several vision-based systems have been designed to identify environmental contexts, pedestrians, obstacles, vehicles, and activities of daily living from video sequences. In particular, convolutional neural networks (CNNs) have shown promising results in place recognition and scene classification in complex indoor and outdoor environments [48].

An increasing number of studies have focused on the automatic analysis of egocentric recordings acquired through wearable cameras and smart glasses. Compared with external viewpoints, these acquisitions provide a more realistic representation of visual exploration and interaction with the environment during movement [49].

Egocentric scene analysis also enable the investigation of contextual factors which influence locomotor behaviour in real-world environments, including pedestrian interactions, environmental complexity, and obstacle negotiation during walking [48].

The integration of wearable scene-recording systems and computer vision algorithms can therefore support a more comprehensive analysis of locomotor behaviour by combining movement-related information with contextual features extracted from the surrounding environment.

When this information is integrated with data acquired from inertial sensors, it becomes possible to analyse simultaneously both the individual's movement and the context in which it occurs. This multimodal approach is particularly relevant in studies of ecological locomotion and pedestrian crossing, where human behaviour is strongly influenced by the characteristics of the surrounding environment.

Chapter 3

Materials and Methods

This chapter describes the materials and methods adopted for the acquisition, annotation, and analysis of the data used in this study. The main objective was to characterise locomotor behaviour and upper-body movements during real-world walking, with particular reference to head dynamics during straight walking and pedestrian crossing situations, distinguishing between conditions with and without an interlocutor.

To achieve this aim, a specific experimental protocol was defined, together with an instrumental configuration based on the combined use of wearable inertial sensors and smart glasses, capable of simultaneously acquiring data related to the kinematics of the body segments under investigation and visual information from the surrounding environment. This approach enabled the investigation of human behaviour in ecologically valid conditions, preserving as much as possible the natural characteristics of walking and of interactions with the urban environment.

Particular attention was devoted to the analysis of behaviour during pedestrian crossings, which were considered an interesting scenario because of the continuous demand for perceptual, motor, and decision-making information. The study involved two cohorts of participants, consisting of younger adults and older adults who were healthy from a psychomotor perspective, in order to assess possible age-related differences in locomotor strategies and behaviour adopted during straight walking and pedestrian crossing.

For this purpose, the acquired data were subjected to synchronisation, recognition, and annotation procedures for the interesting events, allowing the identification of the different experimental conditions analysed.

The following sections describe in detail the instrumentation used, the experimental protocol, the data acquisition and annotation procedures, the processing methods adopted for the extraction of the biomechanical and behavioural parameters analysed in this study, and the statistical analysis procedures used to evaluate the effects of locomotor context, conversation, and age.

The overall methodological workflow adopted in this study is summarised in Figure 3.1.

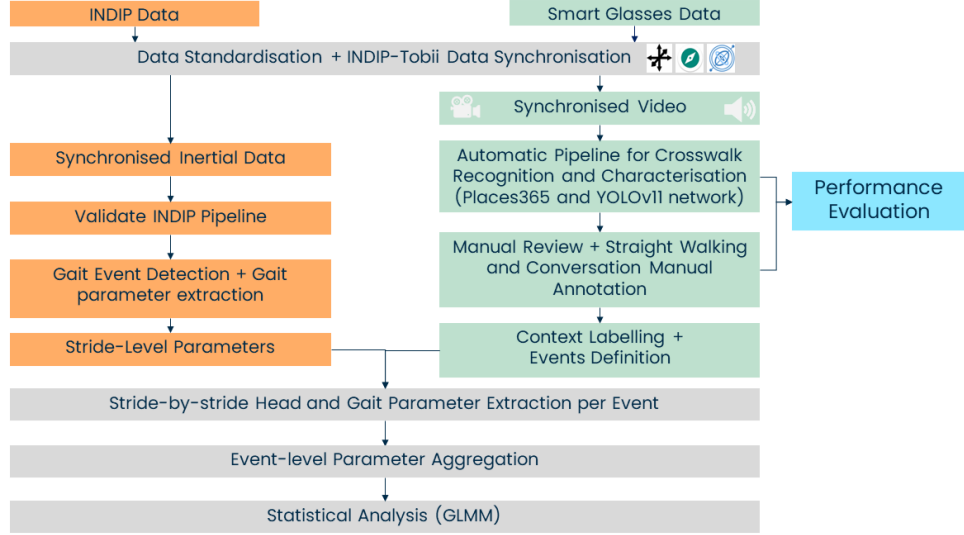


Figure 3.1: Overview of the methodological workflow adopted in the study, from the acquisition of INDIP and smart glasses data to synchronisation, context recognition, manual review, biomechanical parameter extraction, event-level aggregation, and statistical analysis.

3.1 Study Participants

The study involved two cohorts of participants. The first consisted of 10 healthy young adults (4 males and 6 females; age range: 20–31), while the second included 11 healthy older adults (7 males and 4 females; age greater than or equal to 65). All participants reported no neurological, vestibular, or musculoskeletal disorders which could affect postural control and/or locomotion. Participation in the study was voluntary, and all subjects provided written informed consent before the beginning of the experimental activities. The collected data were processed in accordance with the current regulations on the protection of personal data (EU Regulation 2016/679). Due to issues related to data acquisition and/or the subsequent annotation procedures, not all the collected dataset was usable for the final analyses. Therefore, the analyses presented in this work were conducted on a subset of 12 participants, consisting of 6 young adults (3 males and 3 females) and 6 older adults (3 males and 3 females).

Table 3.1: Number of participants recruited and included in the final analyses.

Cohort	Recruited	Included in the analyses
Young adults	10	6
Older adults	11	6
Total	21	12

3.2 Experimental Setup

3.2.1 INDIP System

Inertial data were acquired using the INDIP system (INertial module with DIstance sensors and Pressure insoles), developed at the University of Sassari [50]. INDIP is a wearable multisensor platform designed for the analysis and monitoring of human movement in both controlled laboratory and real-world environments [51, 52, 53, 54]. The fundamental unit of the system consists of a 9-axis Inertial Measurement Unit (IMU), which integrates a tri-axial accelerometer, gyroscope, and magnetometer. Thanks to this configuration, the system enables the measurement of linear accelerations, angular velocities, and body orientation in space. The modular architecture of INDIP also enables integration with additional wearable devices, such as instrumented insoles and distance sensors. The acquired signals were subsequently processed using dedicated algorithms for the identification of gait events and the extraction of the biomechanical parameters of interest. The physical structure of the inertial unit used in the INDIP system is shown in Figure 3.2,

while the main technical specifications of the integrated sensors are summarised in Table 3.2.

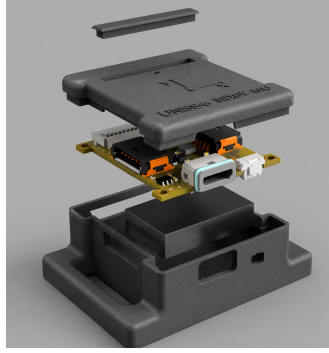


Figure 3.2: Structure of the inertial measurement unit (IMU) used in the INDIP system. The external casing, the electronic board containing the inertial sensors and data acquisition circuits, the integrated battery, and the upper cover of the device are visible.

Table 3.2: Technical specifications of the sensors integrated in the INDIP IMU [50].

Sensor	Parameter	Value	Unit
Tri-axial Accelerometer	Measurement range	± 16	g
	Zero-offset	± 40	mg
	Noise density	1.8–3.0	mg RMS
	Output data rate	1.6–6664	Hz
Tri-axial Gyroscope	Measurement range	± 2000	dps
	Zero-rate offset	± 1	dps
	RMS noise	0.075	dps
	Output data rate	1.6–6664	Hz
Tri-axial Magnetometer	Measurement range	± 50	G
	Zero-G offset	Dynamically cancelled	–
	Noise density	3	mG RMS
	Output data rate	10–100	Hz

3.2.2 Adopted INDIP System Configuration

In the present study, the INDIP system was configured using four Inertial Measurement Units (IMUs), positioned at the head, the lower back, and the two feet. This configuration was chosen to allow the simultaneous assessment of spatiotemporal gait parameters and head movements during locomotion.

- **Head:** an IMU was positioned laterally at the level of the left temporal region using a headband, in order to acquire head movements during walking and to allow temporal synchronisation between the inertial signals from the INDIP system and the data acquired by the smart glasses.
- **Lower back:** an IMU was positioned at the lumbar region at the level of the fifth lumbar vertebra (L5) using an adjustable belt. This sensor was used as a reference for the estimation of locomotor parameters and for the analysis of the relative motion between the head and the trunk.
- **Left foot:** an IMU was fixed to the shoelaces by means of a 3D-printed support. The same unit was connected to an instrumented insole inserted inside the shoe for the measurement of plantar pressures.
- **Right foot:** the IMU was installed on a rigid *L*-shaped support positioned externally on the footwear in the posterior and inner lateral region. Two distance sensors (DS), oriented toward the contralateral foot, were also mounted on the same support and used to measure the distance between the feet and, consequently, to estimate step width. A pressure insole connected to the acquisition system was also used on the right side.

All sensors were synchronised and configured with a sampling frequency of 100 Hz.

The complete configuration of the INDIP system adopted in the study is shown in Figure 3.3.

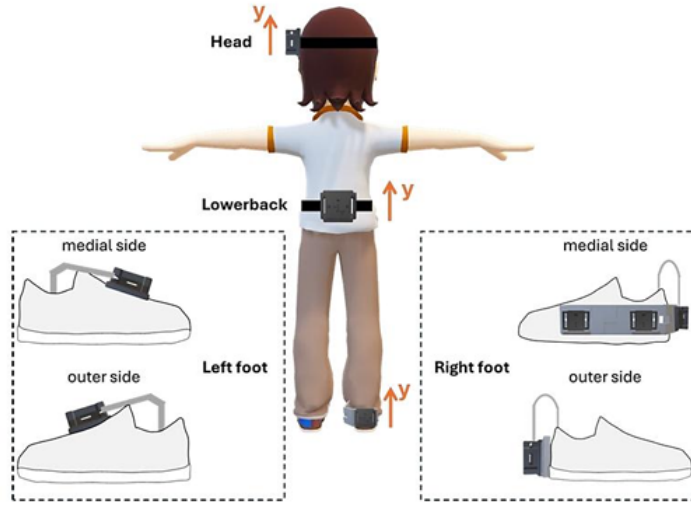


Figure 3.3: Schematic representation of the INDIP system configuration adopted in the study. The inertial sensor positioned on the head, the module installed at the lumbar level, the inertial units mounted on the footwear, and the sensors used to measure the distance between the feet are highlighted.

3.2.3 Smart Glasses for Context Recording

In order to acquire contextual information about the surrounding environment while walking, all participants wore a pair of Tobii Pro Glasses 3 smart glasses. The device is equipped with an integrated front-facing camera, a microphone, and inertial sensors, allowing the simultaneous recording of the scene observed by the participant and the corresponding motion data. The video recordings were used to identify and characterise the contextual events relevant to the study, such as pedestrian crossings with and without traffic lights, face-to-face or phone conversations, and straight walking segments. The information extracted from the videos was subsequently used in the event annotation and classification procedures. The main technical characteristics of the Tobii Pro Glasses 3 used in the study are reported in Table 3.3. Figure 3.4 shows the Tobii Pro Glasses 3 used in the present study.



Figure 3.4: Tobii Pro Glasses 3 used in the present study: device (a) and wearing configuration during data acquisition (b).

Table 3.3: Technical specifications of the Tobii Pro Glasses 3 used in this study, adapted from the manufacturer manual.

Feature	Specification
Model	Tobii Pro Glasses 3
Scene camera resolution	1920 × 1080 pixels at 25 fps
Scene camera video format	H.264
Scene camera field of view	106° diagonal, 95° horizontal, 63° vertical
Eye tracking sampling rate	50 Hz or 100 Hz
Head unit sensors	Gyroscope and accelerometer at 100 Hz; magnetometer at 10 Hz
Audio recording	16-bit mono, integrated microphone
Head unit weight	76.5 g including cable
Recording unit battery life	105 min
Storage media	SD card
Connectors	Micro USB, RJ45, 3.5 mm jack, head unit connector

3.3 Experimental Protocol

In the present study, data acquisition was carried out exclusively in a real-world environment (*Out-of-Lab*), in order to observe the locomotor behaviour of the participants in ecologically valid conditions. Before the beginning of the experimental session, some preliminary procedures were performed, including an initial static phase, a recording in the standing position with the instrumentation worn (*standing trial*), and a *data personalisation* procedure. These acquisitions were used for sensor calibration, the definition of anatomical reference systems, and the initialisation of the processing algorithms employed in the subsequent biomechanical analyses. Subsequently, the participants performed a recording session in a real urban environment while wearing the INDIP system and the smart glasses, enabling the simultaneous recording of inertial data and contextual information.

3.3.1 Out-of-Lab

The *out-of-lab* session was designed to assess the locomotor behaviour of the participants during walking activities performed in a real urban environment, enabling the analysis of gait parameters and head movements in response to natural and dynamic environmental stimuli. This approach was adopted to ensure greater ecological validity compared with traditional assessments conducted in the laboratory. At the end of the preliminary preparation and synchronisation procedures of the instrumentation, the participants completed a predefined walking route near the Politecnico in Turin. In order to place all subjects in similar environmental conditions and behavioural demands, the route was standardised for all participants (Figure 3.5). In addition, at specific points along the route, conversation situations, both in person and on the phone, were introduced thanks to an interlocutor who followed the participant at a distance along the route and who, at predefined positions, approached the participant or made a phone call, in order to subject all participants, as far as possible, to the same attentional and cognitive demands. The route, approximately 3 km long and completed on average in 45 minutes, included straight walking segments used as reference and several pedestrian crossings, both signalised and unsignalised, during which the subject was exposed to different motor, perceptual, and cognitive demands which are typical of daily life.

During the entire session, the participants wore the instrumental configuration described in the previous sections, consisting of the INDIP system and the Tobii Pro Glasses 3. The inertial sensors of the INDIP system continuously recorded locomotor parameters and head movements, while the smart glasses acquired both an egocentric video recording of the environmental context and inertial measurements.

3.4 Data Collection and Processing

3.4.1 Data Acquisition

Data acquisitions were performed using the INDIP system and the Tobii Pro Glasses 3 smart glasses. Before each experimental session, the internal clocks of the INDIP IMU units were synchronised using the INDIP GUI USB software, in order to ensure a common temporal reference across all sensors. Subsequently, the inertial units were connected via Bluetooth Low Energy (BLE) technology to the INDIP App, developed in MATLAB, which was used for device configuration, preliminary functionality checks, and the simultaneous start of the recordings. Once the correct functioning of the smart glasses had been verified, the video recording started and, a few seconds later, the IMU acquisition was launched through the INDIP App.

After both systems had been activated, participants were asked to perform at least three complete head rotations around the vertical axis before the beginning of the route, in order to generate events which could be easily identified in the gyroscopic signals acquired both by the INDIP IMU positioned on the head and by the IMU integrated in the smart glasses. This allowed the subsequent temporal synchronisation between the inertial data from the INDIP system and the video recordings.

During acquisition, the data from the inertial units were stored locally in the internal memory of the devices, enabling participants to move freely without the need to maintain a continuous connection with the computer.

At the end of the route, after again performing three complete head rotations, the recordings were stopped first by interrupting the acquisition of the INDIP system and then that of the Tobii Pro Glasses 3. The IMU data were then downloaded using INDIP GUI USB, while the videos of the scene and the data acquired by the smart glasses were exported through the Tobii Pro Lab software. Finally, all files were organised into specific folders and associated with the participant identifier to facilitate the subsequent phases.

3.4.2 Anthropometric Parameters

In addition to sensor data, several anthropometric parameters were collected for each participant in the study. These parameters were used to personalise the INDIP system processing algorithms and to compute the spatiotemporal gait parameters. The measurements were collected before the beginning of the acquisitions. Table 3.4 reports the list of acquired parameters and a brief description of their meaning.

Table 3.4: Anthropometric measurements collected for each participant and used to personalise the INDIP processing algorithms.

Parameter	Description
Foot length [cm]	Linear distance from the heel to the tip of the longest toe.
Shoe width [cm]	Maximum medio-lateral dimension of the shoe.
Shoe length [cm]	Maximum antero-posterior dimension of the shoe.
Participant age [years]	Age of the participant at the time of testing.
Participant height [cm]	Body height measured before data collection.
Participant weight [kg]	Body weight recorded before the experimental session.
Shoe size [EU]	European shoe size worn during the acquisition.
Posterior waist width [cm]	Trunk width measured at the level where the lower-back sensor was positioned.
Lower back sensor height [cm]	Vertical distance between the lumbar sensor and the ground.
Inter-ASIS distance [cm]	Distance separating the left and right anterior superior iliac spines.
Pelvis-Hip distance [cm]	Distance between the anterior superior iliac spine and the greater trochanter.
Thigh length [cm]	Length of the thigh segment measured from the greater trochanter to the lateral femoral epicondyle.
Shank length [cm]	Length of the shank segment measured from the lateral femoral epicondyle to the lateral malleolus.
Foot height [cm]	Vertical distance between the lateral malleolus and the floor during standing.

3.4.3 Data Standardisation and Synchronisation

Before the extraction of gait and head-movement parameters, the data acquired using the INDIP system and the Tobii Pro Glasses 3 smart glasses were subjected to a series of organisation, standardisation, and synchronisation procedures aimed at creating a unified dataset for the subsequent processing steps.

Data Organisation and Renaming

The first phase of the processing pipeline involved the automatic organisation of the raw files acquired during the experimental sessions. For this purpose, a dedicated MATLAB script was developed and used, capable of automatically identifying the recordings associated with each participant and with the different sensor units of the INDIP system.

The script used the information contained in the *Operator Table*, an Excel table completed by the operator during the experimental session and containing the association between the acquired recordings, the performed exercises, and the sensors used, in order to associate each recording with the corresponding exercise included in the protocol and with the related sensor units involved. Based on this information, the files were automatically renamed according to a standardised nomenclature and organised within a folder structure that was uniform across all participants.

This procedure made it possible to ensure consistent management of the acquisitions, reduce the risk of errors during the subsequent processing steps, and facilitate the automated application of the data standardisation and synchronisation procedures.

INDIP Data Standardisation

During the standardisation phase, the data acquired from the different IMU units of the INDIP system were collected and organised into a unified data structure in MATLAB R2025a. This procedure made it possible to integrate, into a single file, the recordings from the sensors positioned on the head, the lumbar region, and the feet, together with the metadata associated with the participant and the experimental protocol. The anthropometric parameters acquired before the recordings were incorporated into the data structure and used to personalise the processing algorithms. Subsequently, the signals from the different measurement units were subjected to the calibration procedures provided by the INDIP framework and converted into the standard reference system adopted by the project, in accordance with the convention proposed by Palmerini et al. [55]. This procedure ensured a consistent representation of the anatomical axes and inertial measurements across all the sensors used. To standardise the acquisitions from the different IMU units,

the signals were also resampled at a common frequency of 100 Hz and organised according to a hierarchical structure based on participant, experimental session, and acquired trial. The standardised data were finally saved in a MATLAB structure (`data.mat`), which represented the starting point for the subsequent procedures of synchronisation with the smart glasses, gait event identification, and extraction of locomotor parameters.

As shown in Figure 3.6, the INDIP framework uses a common convention for the definition of the anatomical axes and sensor reference systems.

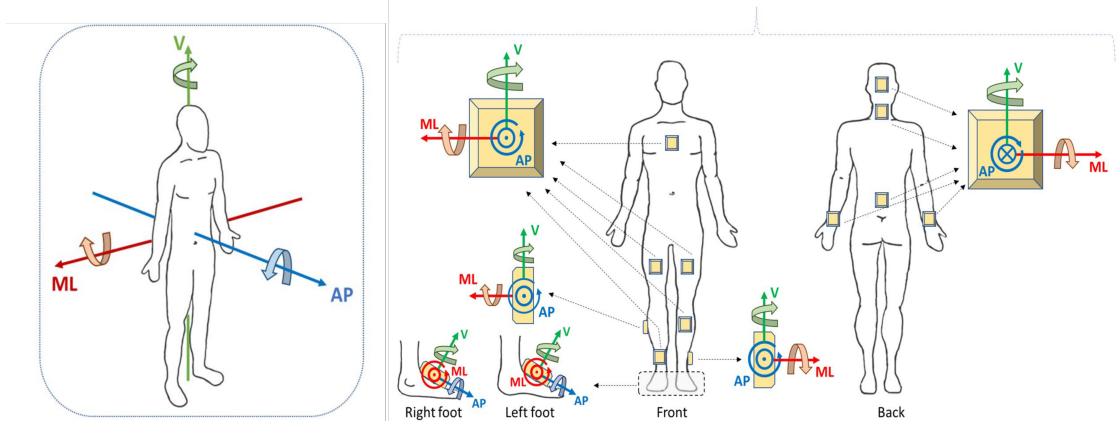


Figure 3.6: Standard reference system adopted by the INDIP framework for the representation of signals acquired from the IMU units. The convention defines the anatomical antero-posterior (AP), medio-lateral (ML), and vertical (V) axes, ensuring a consistent representation of the data from the different sensors. Adapted from Palmerini et al. [55].

Synchronisation between the INDIP System and Tobii Pro Glasses 3

To allow the integration of locomotor information with the data acquired by the smart glasses, a synchronisation procedure between the INDIP system and the Tobii Pro Glasses 3 smart glasses was implemented.

At the end of the acquisitions, the data recorded by the smart glasses were exported using the Tobii Pro Lab software. For each participant, videos of the scene and data from the sensors integrated in the glasses were acquired, including accelerometer, gyroscope, magnetometer, and eye tracker data. The exported data were subsequently imported into MATLAB R2025a and organised into a dedicated structure containing the signals from the different sensors and their corresponding temporal information.

In a first phase, the signals from the different sensors integrated in the smart glasses were resampled onto a common time base. Since these sensors operate

at different sampling frequencies and may present time intervals that do not perfectly overlap, only the temporal period shared by all acquisitions was considered. Subsequently, the reference systems of the Tobii sensors were aligned with those used by the INDIP system through a transformation of the measurement axes, in order to ensure the direct comparability of the inertial signals from the two devices.

Temporal synchronisation was achieved by comparing the gyroscopic signals acquired by the inertial unit positioned on the head within the INDIP system with those recorded by the IMU integrated in the Tobii Pro Glasses 3. For this purpose, the voluntary head movements performed during the acquisition procedure described in Section 3.4.1 were used. These movements generated temporal patterns which could be easily identified in both gyroscope recordings. Starting from the norm of the angular velocity measured by the two devices, the relative temporal delay was automatically estimated through a delay identification procedure.

Figure 3.7 shows an example of the gyroscopic signals acquired by the two systems before and after the application of the synchronisation procedure. Before temporal correction, the patterns associated with the voluntary head movements were misaligned, whereas after delay estimation and compensation, the main oscillations of the two signals overlapped temporally, confirming the correct synchronisation between the acquisitions.

Once the delay had been estimated, the signals from the smart glasses were temporally shifted and subsequently interpolated onto the common time vector used by the INDIP system. This procedure made it possible to obtain a fully synchronised dataset containing inertial signals, eye-tracking data, and video information referred to the same time instant. At the end of the procedure, a unified data structure containing the synchronised signals from the two systems was available for each acquisition, allowing the direct association between gait parameters, head movements, the observed environmental context, and the annotations performed on the egocentric videos. For each participant, the synchronised data were saved in the `data_sync.mat` file. The file contains both the synchronised signals from the two systems and the synchronisation parameters, including the estimated delay and the time instants used for video trimming and for the subsequent creation of the synchronised multimodal dataset.

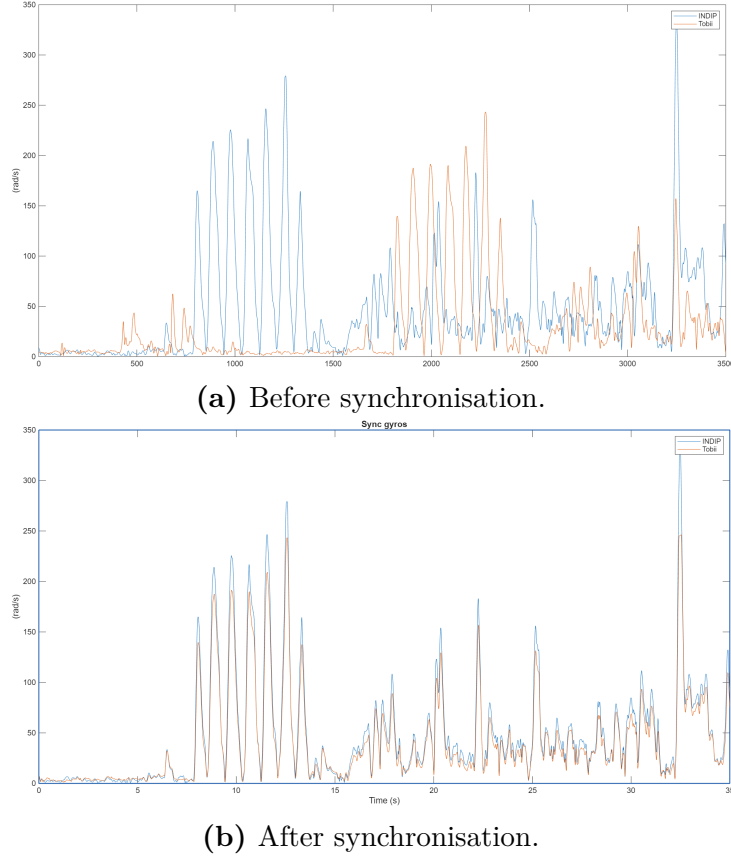


Figure 3.7: Example of the synchronisation procedure between the INDIP system and the Tobii Pro Glasses 3 based on the comparison of the norm of the angular velocity measured by the gyroscopes mounted on the head. In (a), the signals before temporal delay compensation are shown, whereas in (b), the same signals after temporal alignment are reported.

INDIP Pipeline

Once the standardisation and synchronisation phases had been completed, the data were processed using the validated pipeline of the INDIP framework, developed for the automatic identification of gait events and the estimation of the main spatiotemporal parameters [50]. The pipeline jointly uses signals from the IMU units and the instrumented insoles to identify the different phases of gait and segment the recordings into meaningful locomotor sequences.

For each acquisition, the pipeline identified gait sequences, possible changes of direction (*turns*), and individual strides. Based on this information, the main spatiotemporal gait parameters were subsequently computed, including stride length, stride duration, and walking speed. A schematic representation of the

processing pipeline adopted by the INDIP framework is reported in Figure 3.8.

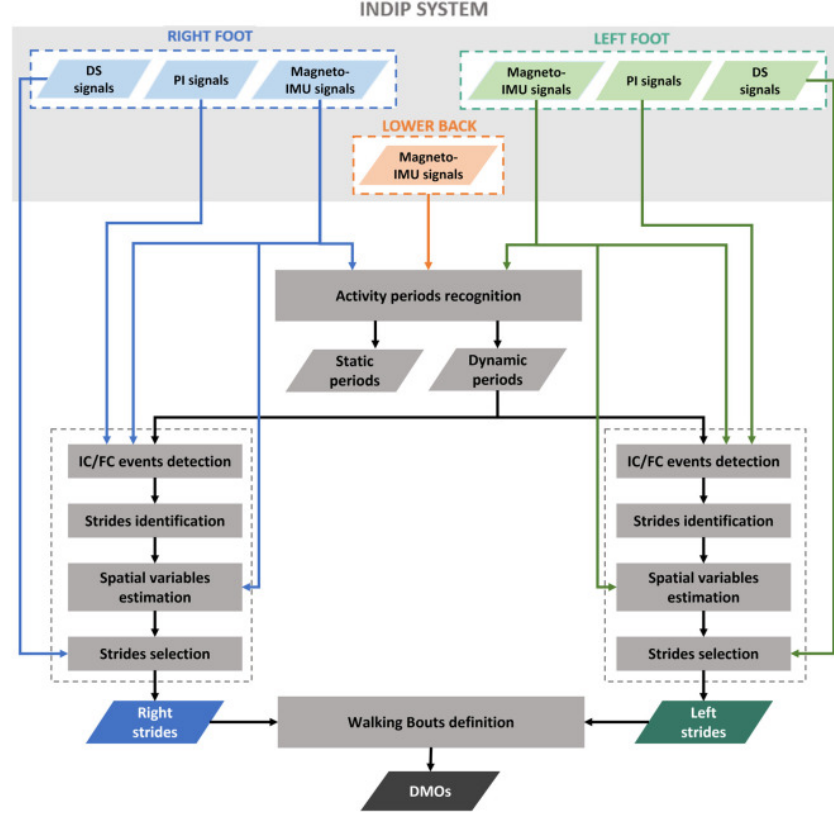


Figure 3.8: Scheme of the processing pipeline used by the INDIP framework for gait event identification and the estimation of spatiotemporal parameters. Adapted from Salis et al. [50].

3.4.4 Video Annotation and Event Segmentation

Video Synchronisation and Cropping

Before event annotation and segmentation, the video recordings acquired using the smart glasses were temporally synchronised with the inertial data from the INDIP system. For this purpose, for each participant, the start (`time_init_INDIP_Tobii`) and end (`time_end_INDIP_Tobii`) instants of the synchronised window, previously identified and saved in the `data_sync.mat` file, were used.

For each subject, the annotated video was loaded, and the acquisition frequency of the video was used to convert the synchronisation time instants into the corresponding frame indices. In particular, the initial and final frames were calculated

from the `time_init_INDIP_Tobii` and `time_end_INDIP_Tobii` times, in order to extract only the portion of video temporally aligned with the INDIP data.

The cropped video sequence was then saved as a new synchronised file for each participant. In addition, a summary table was generated containing, for each subject, the path of the original video, the synchronisation file used, the start and end time instants, the corresponding frames, the number of extracted frames, and the final duration of the synchronised video. This procedure made it possible to obtain videos temporally consistent with the inertial signals acquired during the protocol for all the subsequent phases of context recognition and event segmentation.

Scene Recognition and Temporal Stabilisation

The synchronised video recordings were processed in order to automatically identify the pedestrian crossings present along the experimental route. For this purpose, the pre-trained Places365-GoogLeNet network [56] was used, without any additional training or fine-tuning procedures.

Although Places365 is able to recognise hundreds of different environmental categories, in the present study the analysis was limited to the *crosswalk* category, corresponding to pedestrian crossings. This choice made it possible to focus the classification process on the elements of interest for the subsequent segmentation of the route, reducing the possibility of assignments to non-relevant contexts. To reduce the computational load while maintaining adequate temporal resolution, one frame every five was processed; the obtained predictions were then extended to the intermediate non-processed frames, allowing a continuous representation of the entire video sequence.

For each processed frame, the network provided a probability of belonging to the *crosswalk* category. The temporal probability sequences obtained were then subjected to a temporal stabilisation procedure. In a first step, an Exponential Moving Average (EMA) was applied in order to reduce short-term fluctuations in the network predictions. The EMA was computed recursively according to the following equation:

$$P_{EMA}(t) = \alpha P(t) + (1 - \alpha)P_{EMA}(t - 1) \quad (3.1)$$

where $P(t)$ represents the probability of belonging to the *crosswalk* category provided by the network for the current frame, $P_{EMA}(t)$ is the filtered value, and α is the smoothing coefficient ranging between 0 and 1. Higher values of α assign greater weight to the most recent observations, whereas lower values produce stronger attenuation of temporal fluctuations.

Subsequently, the probabilities filtered using the EMA were further smoothed through a moving average computed over a temporal window of 0.4 s, defined

according to the effective frequency of the processed frames. Hysteresis criteria were then applied to the resulting probabilities to determine the activation and deactivation of the recognised events. Finally, events shorter than the predefined duration thresholds were automatically removed, reducing the influence of isolated classifications and spurious transitions. The main parameters used in the temporal stabilisation procedure are reported in Table 3.5.

This procedure made it possible to obtain a temporal sequence of labels that was more stable and consistent with the actual evolution of the context observed during walking. The final output was used in the subsequent phases of the event annotation and segmentation pipeline.

Table 3.5: Parameters used for the temporal stabilisation of Places365 predictions.

Parameter	Value	Description
Frame sampling	1 every 5 frames	Frame processing frequency
EMA coefficient (α)	0.30	Exponential moving average coefficient
Moving average window	0.4 s	Temporal filtering window
Crosswalk ON threshold	0.20	Event activation threshold
Crosswalk OFF threshold	0.12	Event deactivation threshold
Minimum event duration	0.6 s	Minimum duration of the recognised event

Traffic Light Detection

The identification of traffic lights in the video recordings was performed using YOLOv11 [57], a neural network for object detection. The algorithm was applied exclusively to frames belonging to segments previously classified as pedestrian crossings (*crosswalks*), reducing the number of non-relevant processing operations and the overall computational load.

For each analysed frame, the network provided the possible position of traffic lights present in the image through bounding boxes associated with a confidence value. Since multiple traffic lights could be present simultaneously in some frames, a selection procedure was implemented to identify the traffic lights which were the most relevant to the participant. This selection was performed by considering the detection confidence and the size of the bounding box of the identified region.

To ensure greater robustness of identification over time, a tracking mechanism based on the overlap between successive bounding boxes and on the distance between the positions detected in consecutive frames was also implemented. This approach made it possible to maintain the association with the same traffic lights

even in the presence of brief occlusions or temporary detection losses.

Once the traffic lights of interest had been identified, the corresponding region of interest was used to estimate the state of the light signal. The procedure adopted for traffic light state classification is described in Section ‘Traffic Light State Classification’. The main parameters used for traffic light detection, tracking, and temporal stabilisation are reported in Table 3.6.

At the end of the procedure, for each frame, the possible region of interest associated with the traffic lights which were relevant to the participant was available. In cases where no detectable traffic light was present, the segment was considered as not being controlled by traffic lights.

Although the pipeline allowed the identification of the traffic lights observed by the participant, a qualitative inspection of the results highlighted some limitations in the robustness of the association with the individual traffic light device. Since a quantitative evaluation of this aspect is beyond the objectives of the present work and will be addressed in future analyses, the output of the YOLOv11-based module was used exclusively to classify each crossing episode as either a signalised crossing (*Signalised Crossing*) or an unsignalised crossing (*Unsignalised Crossing*).

Table 3.6: Parameters used for traffic light detection and tracking.

Parameter	Value	Description
YOLO confidence threshold	0.25	Minimum confidence required to accept a detection produced by YOLOv11.
Tracking IoU threshold	0.20	Minimum overlap between consecutive bounding boxes required to maintain the association with the same traffic lights.
Maximum missed detections	12 frames	Maximum number of consecutive frames without detection before interrupting the traffic light tracking.

Traffic Light State Classification

Once the position of the traffic lights within the image had been identified, a dedicated procedure was performed to determine its light state. Unlike traffic light detection, which was based on the YOLOv11 network, colour classification was performed through an analysis of the region of interest (ROI) corresponding to the detected bounding box.

For each frame, the region containing the traffic lights was extracted and converted into the HSV colour space (*Hue, Saturation, Value*), which is particularly suitable for separating chromatic information from information related to light intensity.

Since the traffic lights along the route were characterised by a vertical arrangement of the lamps, the region of interest was divided into three portions corresponding to the approximate positions of the red, amber, and green lights. The combined analysis of the brightness distribution and chromatic components allowed the active lamp within the region of interest to be identified and one of the following states to be assigned to the frame: *red*, *amber*, *green*, or *unknown*.

The adopted approach made it possible to treat traffic light detection and classification of its light state separately. The chromatic analysis was therefore performed exclusively within the regions previously identified as traffic lights, reducing the influence of elements present in the rest of the scene and allowing the automatic identification of the different phases of the light signal.

Since lighting conditions, temporary occlusions, and the environmental variability typical of real-world acquisitions may generate unstable classifications, a temporal stabilisation procedure for the traffic light state was implemented. In particular, transitions between consecutive states were confirmed only after being observed in multiple consecutive frames, whereas brief absences of reliable classifications were handled by temporarily maintaining the last available valid state. In addition, the observed transitions were subjected to temporal validation criteria in order to reduce the influence of spurious classifications. The main parameters used for traffic light state classification and temporal stabilisation are reported in Table 3.7.

Despite these precautions, a qualitative inspection of the results highlighted that the high variability in the shapes, apparent sizes, and visual characteristics of the traffic lights present in the recordings made the automatic approach not sufficiently robust to guarantee reliable traffic light state labelling across all recordings. Therefore, the automatic classification of the light state was not used in the final analyses, and the labelling of the traffic light phases (*red*, *amber*, and *green*) was subsequently performed through manual segmentation.

Table 3.7: Parameters used for traffic light state classification and temporal stabilisation.

Parameter	Value
colour confidence threshold	0.45
Temporal hold duration	1.0 s
State-change confirmation frames	4 frames

Manual Review and Conversation Annotation

The results obtained from the automatic procedures for context recognition, traffic light detection, and traffic light state classification were subsequently subjected to a manual review phase. This phase was introduced to verify the correctness of the

automatically generated labels and to correct possible errors due to unfavourable lighting conditions, temporary occlusions, changes in the field of view, the high variability of traffic light configurations in the urban environment, and limitations of the recognition algorithms.

The review was performed using dedicated interfaces developed in MATLAB, which allowed the synchronised visualisation of the video recordings and of the labels generated by the automatic pipeline. During this phase, it was possible to correct the automatically identified contexts, verify the presence of traffic lights and manually annotate their light state, as well as associate additional information which could not be obtained through the automatic procedures. Particular attention was devoted to the identification of conversation conditions. For each event, the presence or absence of in person conversations with other people and phone conversations was annotated. Although phone-related interactions were included during the annotation phase, they were not considered in the final statistical analyses. The collected information was subsequently used to classify the events into the *Conversation* and *No Conversation* conditions.

At the end of the review, a CSV file containing the final temporal labels associated with the different events observed along the route was generated for each participant. These annotations were subsequently used to segment the synchronised data from the INDIP system and the smart glasses, to associate locomotor parameters and head movements with the corresponding contexts analysed in the study, and as *ground truth* for the evaluation of the performance of the automatic procedures for context recognition and classification of crossings into signalised and unsignalised contexts (Section 3.5). The same annotations were also used for the extraction and aggregation of the biomechanical parameters of gait and head movements described in Section 3.6.

Final Event Annotation Structure

The final annotations generated at the end of the manual review consisted of two levels of information: the locomotor context and the conversation condition.

The locomotor context identified the type of event observed during the experimental route, whereas the conversation condition described the presence or absence of conversation associated with the event. The categories used for the annotation of locomotor contexts are reported in Table 3.8, while the annotated conversation conditions are reported in Table 3.9.

Table 3.8: Categories used for event annotation.

Category	Description
TR	Straight walking
CWTLG	Crossing with a green traffic light
CWTLY	Crossing with an amber traffic light
CWTLR	Crossing with a red traffic light
CWNOTL	Crossing without traffic lights
CWTLunknown	Unidentified traffic light state
UNDEF	Undefined context

Table 3.9: Conversation categories annotated during the manual review.

Category	Description
No Conversation	No conversation observed
Conversation	Face-to-face conversation with another person
Phone Conversation	Conversation on mobile phone

For the subsequent statistical analyses, pedestrian crossings performed with a green traffic light (*Green Traffic Light*) or an amber traffic light (*Amber Traffic Light*) were aggregated into the *Signalised Crossing* context, whereas crossings without traffic lights (*Crossing without Traffic Light*) were classified as *Unsignalised Crossing*. Crossings performed with a red traffic light (*Red Traffic Light*) and those characterised by an unidentified traffic light state (*Unknown Traffic Lights*) were not included in the final analyses. Similarly, events associated with *Phone Conversation* were excluded from the subsequent statistical processing.

The combination of locomotor context and conversation condition allowed the definition of six analysis conditions: *Straight Walking No Conversation*, *Straight Walking Conversation*, *Signalised Crossing No Conversation*, *Signalised Crossing Conversation*, *Unsignalised Crossing No Conversation*, and *Unsignalised Crossing Conversation*.

3.5 Evaluation of the Performance of the Automatic Pipeline

The performance of the automatic pipeline for pedestrian crossing and traffic light recognition was evaluated by comparing the automatically generated labels with the manual annotations, which were considered as *ground truth*. The evaluation was performed both at frame level and at event level, in order to quantify the contribution of post-processing in terms of classification accuracy and reduction of spurious detections produced by the raw output of the classifier.

The automatic pipeline included two main modules: a pedestrian crossing recognition module based on the Places365 classifier and a traffic light detection module based on the YOLOv11 object detection model. Both modules initially produce an output at the single-frame level, which is subsequently used to generate temporally coherent events and to associate contextual information with the biomechanical acquisitions.

For pedestrian crossing recognition, performance was evaluated both on the raw frame-level output (*RAW*), produced directly by the Places365 classifier, and on the output obtained after applying the temporal post-processing procedures developed in the present study (*PROCESSED*). The comparison between the two configurations was performed using the same manual annotations as *ground truth*, in order to quantify the contribution of post-processing in terms of classification accuracy and reduction of spurious detections produced by the raw output of the classifier.

For traffic light detection, the performance of the YOLOv11 model was evaluated by comparing the frames in which the presence of a traffic lights were automatically detected with the corresponding manual annotations. However, since the final objective of the study was not the simple identification of traffic lights in individual frames, but rather the characterisation of the type of pedestrian crossing, distinguishing between crossings regulated by traffic lights and crossings not regulated by traffic lights, an event-level evaluation was also performed by classifying each crossing episode as either a signalised crossing or an unsignalised crossing.

3.5.1 Evaluation of Pedestrian Crossing Recognition Performance

Frame-Level Evaluation

For each video frame, the algorithm produced a binary classification between the *Crosswalk* and *Non-Crosswalk* classes. The automatically generated labels were compared with the manual annotations, defined as *ground truth*.

The following quantities were defined:

- **True Positive (TP):** frames classified as a crossing both automatically and manually;
- **True Negative (TN):** frames classified as non-crossing both automatically and manually;
- **False Positive (FP):** frames incorrectly classified as a crossing by the automatic procedure;
- **False Negative (FN):** frames belonging to an actual crossing but not automatically detected.

Starting from these quantities, the following metrics were calculated:

Accuracy: Represents the proportion of correctly classified frames compared to the total number of analysed frames and it provides an overall measure of classifier performance.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.2)$$

Precision: Measures the proportion of frames classified as crossings which actually belong to a crossing in the *ground truth*. It therefore provides a measure of the reliability of positive detections.

$$Precision = \frac{TP}{TP + FP} \quad (3.3)$$

Recall: Measures the ability of the algorithm to correctly identify all frames belonging to an actual crossing and therefore represents a measure of the sensitivity of the classifier.

$$Recall = \frac{TP}{TP + FN} \quad (3.4)$$

F1-score: Combines precision and recall into a single measure, providing an overall indicator of classifier performance.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (3.5)$$

The frame-level analysis was used to quantify the performance of the classification of individual frames independently of their temporal organisation.

Event-Level Evaluation

Since the final objective of the pipeline was not exclusively the classification of individual frames, but rather the identification of crossing events which could be used in the subsequent biomechanical analyses, an additional event-level evaluation was performed. This analysis was introduced in particular to assess the effectiveness of the temporal post-processing procedures and the correspondence between the automatically identified crossing events and those manually annotated.

Consecutive frames classified as *Crosswalk* were grouped into single events for both the manual annotations and the automatic output.

For each pair of manual and automatic events, their temporal intersection was calculated. Two temporal overlap metrics were therefore defined:

$$Overlap_{manual} = \frac{N_{intersection}}{N_{manual}} \quad (3.6)$$

$$Overlap_{automatic} = \frac{N_{intersection}}{N_{automatic}} \quad (3.7)$$

where:

- $N_{intersection}$ represents the number of frames belonging to the intersection between the two events;
- N_{manual} represents the duration of the manual event;
- $N_{automatic}$ represents the duration of the automatic event.

A manual event was considered correctly detected when at least one automatic event showed a temporal intersection exceeding a minimum threshold expressed as a percentage of both the duration of the manual event and the duration of the automatic event. A bidirectional overlap threshold of 20% was used as the main evaluation criterion. In addition, in order to assess the robustness of the results compared to the choice of the threshold, a sensitivity analysis was performed considering overlap values of 10%, 20%, 30%, 50%, and 70%.

For the event-level evaluation, the following quantities were defined:

- **Manual events:** total number of manually annotated crossing events;
- **Predicted events:** total number of crossing events automatically generated by the algorithm;
- **Detected events:** number of manual events for which at least one corresponding automatic event was identified;

- **Missed events:** number of manual events for which no corresponding automatic event was identified;
- **Extra events:** number of automatic events which did not show any correspondence with a manual event;
- **Fragmentation index:** number of additional automatic fragments associated with already detected manual events.

Since a single manually annotated event could be associated with multiple automatic events, the number of predicted events did not necessarily correspond to the sum of detected events and extra events. This discrepancy was quantified using a fragmentation index, defined as:

$$FI = N_{\text{matched predicted}} - N_{\text{detected}} \quad (3.8)$$

where $N_{\text{matched predicted}}$ represents the number of predicted events associated with at least one manual event, whereas N_{detected} represents the number of manual events detected by at least one predicted event. Consequently, the total number of predicted events can be expressed as:

$$\text{PredictedEvents} = \text{DetectedEvents} + \text{ExtraEvents} + FI \quad (3.9)$$

Starting from these quantities, event-level recall, precision, and F1-score were calculated:

$$\text{Recall}_{\text{event}} = \frac{\text{DetectedEvents}}{\text{ManualEvents}} \quad (3.10)$$

$$\text{Precision}_{\text{event}} = \frac{\text{DetectedEvents}}{\text{DetectedEvents} + \text{ExtraEvents}} \quad (3.11)$$

$$F1_{\text{event}} = \frac{2 \cdot \text{Precision}_{\text{event}} \cdot \text{Recall}_{\text{event}}}{\text{Precision}_{\text{event}} + \text{Recall}_{\text{event}}} \quad (3.12)$$

where ManualEvents represents the total number of manually annotated events. Event-level precision was calculated by considering extra events as false positives, whereas fragmentation was reported separately to describe the presence of multiple automatic fragments associated with the same manual event.

Table 3.10: Metrics used to evaluate the performance of pedestrian crossing recognition.

Metric	Level	Description
Accuracy	Frame	Proportion of correctly classified frames compared to the total number of analysed frames.
Precision	Frame/Event	Reliability of positive detections, defined as the proportion of automatically detected frames or events which correspond to the <i>ground truth</i> .
Recall	Frame/Event	Ability of the algorithm to correctly identify the frames or events present in the <i>ground truth</i> .
F1-score	Frame/Event	Harmonic mean between precision and recall, used as an overall measure of classification performance.
Manual events	Event	Total number of manually annotated crossing events.
Predicted events	Event	Total number of crossing events automatically generated by the algorithm.
Detected events	Event	Number of manual events for which at least one corresponding automatic event was identified.
Missed events	Event	Number of manual events not detected by the automatic procedure.
Extra events	Event	Number of automatic events which did not show any correspondence with manual events.
Fragmentation index	Event	Number of additional automatic fragments associated with already detected manual events.

3.5.2 Evaluation of Traffic Light Recognition Performance

Frame-Level Evaluation

The performance of the traffic light detection algorithm was evaluated at frame level by comparing the automatically generated labels with the manual annotations, considered as *ground truth*.

For each frame, a binary classification was performed between the *Traffic Light* and *No Traffic Light* classes.

The following quantities were defined:

- **True Positive (TP):** frames in which the presence of a traffic light were

identified both automatically and manually;

- **True Negative (TN):** frames in which the traffic lights were absent both according to the algorithm and the manual annotation;
- **False Positive (FP):** frames in which a traffic lights were automatically detected but they were not present in the *ground truth*;
- **False Negative (FN):** frames in which a traffic lights were present in the *ground truth* but they were not automatically detected.

Starting from these quantities, the following metrics were calculated:

Accuracy: Represents the proportion of correctly classified frames compared to the total number of analysed frames and provides an overall measure of classifier performance.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.13)$$

Precision: Measures the proportion of frames in which the traffic lights were automatically detected and that actually correspond to the presence of traffic lights in the *ground truth*.

$$Precision = \frac{TP}{TP + FP} \quad (3.14)$$

Recall: Measures the ability of the algorithm to correctly identify all frames containing traffic lights and therefore represents a measure of classifier sensitivity.

$$Recall = \frac{TP}{TP + FN} \quad (3.15)$$

Specificity: Measures the ability of the algorithm to correctly identify frames in which the traffic lights are absent and therefore represents the true negative rate.

$$Specificity = \frac{TN}{TN + FP} \quad (3.16)$$

F1-score: Combines precision and recall into a single measure, providing an overall indicator of classifier performance.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (3.17)$$

The frame-level analysis was used to evaluate the ability of the algorithm to correctly identify the presence of traffic lights in individual frames, independently of the subsequent classification of crossing events.

Event-Level Evaluation

Since the final objective of the study was to characterise pedestrian crossings by distinguishing between crossings regulated by traffic lights and crossings not regulated by traffic lights, an event-level evaluation was also performed.

Crossing events were defined based on the manual annotations, considered as *ground truth*. In particular, both crossings regulated by traffic lights and crossings without traffic lights were identified.

For each manually annotated event, the percentage of frames in which the traffic lights were automatically detected was then calculated:

$$P_{TL} = \frac{N_{TL}}{N_{event}} \cdot 100 \quad (3.18)$$

where:

- N_{TL} represents the number of frames of the event in which the traffic lights were automatically detected;
- N_{event} represents the total number of frames belonging to the crossing event.

A crossing event was classified as a *Signalised Crossing* when the traffic lights were detected in at least 20% of the frames belonging to the event.

In addition, in order to evaluate the robustness of the results compared to the choice of the classification threshold, a sensitivity analysis was performed by considering thresholds of 10%, 20%, 30%, 40%, 50%, and 60%.

A threshold lower than 50% was used because, during a real crossing, the traffic lights do not necessarily remain visible for the entire duration of the event. Indeed, the device records the scene from the participant's point of view, and the traffic lights may temporarily leave the field of view due to natural head movements, gaze orientation, or environmental occlusions.

Based on these considerations, a threshold of 20% was used as the main criterion for event classification.

For the events manually annotated as crossings with traffic lights, descriptive statistics of the percentage of frames classified as containing traffic lights were also calculated separately for correctly classified events (True Positive) and for events not automatically identified (False Negative).

The following quantities were therefore defined:

- **True Positive (TP):** correctly identified crossing with traffic light;

- **True Negative (TN):** correctly classified crossing without traffic lights;
- **False Positive (FP):** crossing without traffic lights incorrectly classified as regulated by traffic lights;
- **False Negative (FN):** crossing with traffic lights not automatically identified.

The metrics of Accuracy, Precision, Recall, Specificity, and F1-score were also subsequently calculated at event level using the definitions reported above.

3.6 Extraction of Biomechanical Parameters of Gait and Head Movement

The quantitative analysis of locomotor behaviour was conducted by extracting spatiotemporal gait parameters and descriptive measures of head movement. Gait parameters were used to characterise locomotor performance in the different environmental contexts, while the signals acquired from the inertial unit positioned on the head were used to quantify different aspects of head motor control.

In particular, movement amplitude parameters were calculated as the root mean square of acceleration normalised by walking speed (*speed-normalised acceleration root mean square*, aRMS/speed), regularity parameters using the *Refined Composite Multiscale Entropy* (RCME), and smoothness parameters using the *Log Dimensionless Jerk* (LDLJ). All head-movement parameters were extracted along the vertical (V), medio-lateral (ML), and antero-posterior (AP) directions.

The obtained parameters were subsequently associated with the contextual information derived from the analysis of the egocentric videos and used to evaluate the effects of age, locomotor context, and the presence of conversation during urban locomotion.

To allow the association between biomechanical parameters and the different locomotor contexts, the annotations obtained from the review of the egocentric videos were synchronised with the inertial signals from the INDIP system. This procedure made it possible to segment the data according to the events of interest and to assign the corresponding locomotor context and conversation condition to each time interval. Figure 3.9 reports an example of a plot used to visually verify the correct temporal correspondence between pedestrian crossing annotations and the medio-lateral gyroscopic signals acquired from the sensors positioned on the feet.

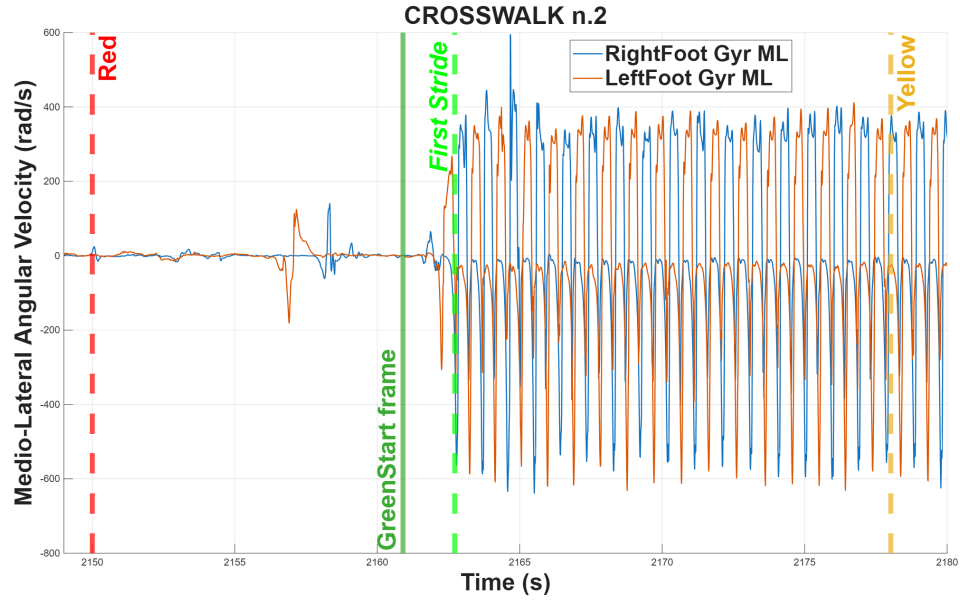


Figure 3.9: Example of temporal alignment between event annotations obtained from egocentric videos and inertial signals acquired by the INDIP system. The figure shows the correspondence between pedestrian crossing annotations and the medio-lateral gyroscopic signals recorded by the sensors positioned on the feet, used to verify the correct synchronisation between contextual information and biomechanical data.

Table 3.11 summarises the extracted biomechanical parameters and the corresponding aspect of locomotor behaviour they describe.

Table 3.11: Extracted biomechanical parameters and corresponding physiological meaning.

Parameter	Category	Meaning
Froude Number	Gait	Normalised speed
Stride Speed	Gait	Locomotor speed
Stride Length	Gait	Stride length
Stride Duration	Gait	Stride duration
Cadence	Gait	Step frequency
Speed-normalised aRMS	Head	Movement amplitude
RCME	Head	Regularity/complexity
LDLJ	Head	Movement smoothness

3.6.1 Preprocessing: Orientation Estimation and Gravity Removal

Before the extraction of biomechanical parameters, the inertial signals acquired from the units positioned on the head and on the lower back were preprocessed in order to estimate their orientation and remove the contribution of gravity.

Orientation estimation was performed using the processing pipeline developed within the INDIP project, based on the fusion of information from the accelerometer, gyroscope, and magnetometer, according to a procedure described in previous works [58]. This procedure made it possible to obtain, for each time instant, the quaternions representing the orientation of the sensors compared to the global reference system.

The estimated quaternions were subsequently used to rotate the accelerometric signals into the global reference system and remove the gravitational component. The resulting accelerations were then expressed in the lower-back reference system, allowing direct comparison between head and trunk movements and the extraction of the biomechanical parameters described in the following sections.

3.6.2 Spatiotemporal Gait Parameters

Spatiotemporal gait parameters were extracted using the algorithms developed within the INDIP project, applied to the signals acquired from the inertial units positioned on the feet. Gait events were automatically identified and used to compute the main descriptive parameters of locomotion.

In particular, for each stride, walking speed (*stride speed*), stride length (*stride length*), stride duration (*stride duration*), and cadence (*cadence*) were calculated.

Froude Number

In order to compare walking speed among subjects characterised by different anthropometric features, the Froude Number (FN) was calculated. This dimensionless parameter is commonly used in locomotion studies to normalise speed compared to body dimensions [59].

The Froude Number is defined as:

$$FN = \frac{v^2}{gL} \quad (3.19)$$

where v represents walking speed, g gravitational acceleration, and L a characteristic length of the subject.

In the present study, the characteristic length L was assumed to be equal to the total length of the lower limb. This normalisation allows locomotor performance

to be compared among individuals of different stature, reducing the influence of anthropometric differences on walking speed.

For each locomotor event, the Froude Number value was calculated using the mean walking speed within the event and it was subsequently used in the statistical analyses.

3.6.3 Head-Movement Parameters

In addition to spatiotemporal gait parameters, several descriptive measures of head movement were extracted from the accelerometric signals acquired by the inertial unit positioned on the head. These parameters were selected to characterise different aspects of motor control during locomotion, including the amplitude of head oscillations, the temporal regularity of the movement, and its smoothness.

All analyses were performed separately along the vertical (V), medio-lateral (ML), and antero-posterior (AP) axes, using the linear accelerations obtained after sensor orientation estimation and removal of the gravitational component.

Speed-Normalised Acceleration Root Mean Square

The amplitude of head accelerations when walking provides quantitative information on the intensity of head movement and represents a relevant aspect for the characterisation of locomotor control.

To quantify the magnitude of head accelerations, the acceleration Root Mean Square (aRMS) was used, a measure which summarises the mean intensity of the accelerations recorded during movement.

For each anatomical axis, aRMS was calculated as:

$$aRMS = \sqrt{\frac{1}{N} \sum_{k=1}^N a(k)^2} \quad (3.20)$$

where $a(k)$ represents the acceleration value at sample k , and N is the total number of samples belonging to the analysed event.

Higher aRMS values indicate accelerations of greater amplitude, whereas lower values reflect accelerations of smaller magnitude.

Since the amplitude of head accelerations is influenced by locomotion speed, the parameter was normalised with respect to the mean walking speed of the event:

$$aRMS/speed = \frac{aRMS}{v} \quad (3.21)$$

where v represents the mean walking speed. This procedure made it possible to compare different subjects and conditions by reducing the effect of variations

in locomotor speed and allowing a more direct comparison of head-movement characteristics across subjects and experimental conditions.

Regularity of Head Movement

The regularity of head movement was evaluated using *Refined Composite Multiscale Entropy* (RCME), a nonlinear measure proposed by Wu et al. [60] to quantify the complexity of a time series at different temporal scales. Multiscale entropy methods make it possible to evaluate how the complexity of a signal changes as its temporal resolution varies.

In the present study, RCME was calculated using the MATLAB code implemented by Ihlen et al. [61] and applied to the linear accelerations of the head along the vertical (V), medio-lateral (ML), and antero-posterior (AP) directions. To reduce the influence of differences in event duration on the estimation of the parameter, all accelerometric signals were resampled to 2000 samples using *Radial Basis Function* (RBF) interpolation.

The first step in the calculation of RCME consists of constructing the *coarse-grained* time series. For a scale factor τ , the original signal $x(i)$ is divided into non-overlapping windows of length τ , and the samples contained in each window are averaged. The k -th *coarse-grained* time series is therefore defined as:

$$y_k^{(\tau)}(j) = \frac{1}{\tau} \sum_{i=(j-1)\tau+k}^{j\tau+k-1} x(i) \quad (3.22)$$

where $y_k^{(\tau)}(j)$ represents the j -th sample of the k -th *coarse-grained* time series, τ is the scale factor, and k is the index of the reconstructed series.

For each *coarse-grained* time series, template vectors of dimension m are subsequently constructed:

$$\mathbf{y}_{k,\tau}^m(j) = [y_k^{(\tau)}(j), y_k^{(\tau)}(j+1), \dots, y_k^{(\tau)}(j+m-1)]. \quad (3.23)$$

Two template vectors are considered similar when the distance between them is lower than a tolerance threshold r . In the present study, template vectors of dimension

$$m = 2 \quad (3.24)$$

and a tolerance equal to

$$r = 0.1 \cdot \text{std}(\mathbf{x}) \quad (3.25)$$

were used, where $\text{std}(\mathbf{x})$ represents the standard deviation of the analysed signal.

For each scale factor, RCME is calculated as the negative logarithm of the ratio between the mean number of pairs of similar vectors of length $m + 1$ and the mean number of pairs of similar vectors of length m :

$$RCME(\mathbf{x}, \tau, m, r) = -\ln \left(\frac{\bar{n}_{k,\tau}^{m+1}}{\bar{n}_{k,\tau}^m} \right) = -\ln \left(\frac{\sum_{k=1}^{\tau} n_{k,\tau}^{m+1}}{\sum_{k=1}^{\tau} n_{k,\tau}^m} \right) \quad (3.26)$$

where $\bar{n}_{k,\tau}^m$ and $\bar{n}_{k,\tau}^{m+1}$ represent, respectively, the mean number of pairs of similar vectors of length m and $m + 1$ calculated over all the coarse-grained time series.

In the present study, scale factors ranging from $\tau = 1$ to $\tau = 6$ were considered. To obtain a single value representative of signal complexity, the overall entropy, hereafter simply referred to as RCME, was calculated as:

$$RCME(\mathbf{x}, m, r) = \sum_{\tau=1}^6 RCME(\mathbf{x}, \tau, m, r). \quad (3.27)$$

Smoothness of Head Movement

The smoothness of head movement was evaluated using the *Log Dimensionless Jerk* (LDLJ), an index widely used to quantify movement fluidity and the continuity of motor control [62, 63]. Smoothness represents an important indicator of the quality of motor control and describes the continuity of movement and the absence of abrupt variations over time.

LDLJ is based on the concept of *jerk*, defined as the time derivative of acceleration:

$$j(t) = \frac{da(t)}{dt} \quad (3.28)$$

Since jerk depends on both the amplitude and the duration of the movement, its direct use does not allow comparison between different events. To obtain a measure independent of signal scale, a dimensionless formulation of jerk was therefore introduced [62].

In applications based on optoelectronic systems, LDLJ is generally calculated from the movement velocity profile. However, in applications based on inertial sensors, obtaining velocity requires integration procedures which may introduce errors due to signal noise and drift. For this reason, an alternative formulation based directly on acceleration was adopted, which is particularly suitable for the analysis of data from wearable inertial sensors [63].

LDLJ was therefore calculated as:

$$LDLJ = -\ln \left(\frac{t_2 - t_1}{a_{peak}^2} \int_{t_1}^{t_2} \left| \frac{da(t)}{dt} \right|^2 dt \right) \quad (3.29)$$

where t_1 and t_2 represent the beginning and the end of the analysed interval, respectively, while a_{peak} represents the maximum acceleration value recorded within the event.

The parameter was calculated separately along the vertical (V), medio-lateral (ML), and antero-posterior (AP) axes using the linear accelerations of the head obtained after removal of the gravitational component.

More negative LDLJ values indicate lower movement smoothness, whereas less negative values reflect smoother and more continuous movements.

3.7 Obtained Dataset

The original dataset consisted of stride-level observations obtained from the integration of the biomechanical parameters extracted from the INDIP system and the contextual annotations derived from the analysis of the egocentric videos acquired using smart glasses. For each stride, biomechanical parameters related to gait and head movement were available, together with information regarding the subject, the locomotor context, and the presence of a conversation.

Before the construction of the final dataset, an observation selection phase was performed in order to retain only the data consistent with the objectives of the study. In particular, strides belonging to turning phases were excluded, as they were characterised by different motor demands compared with straight locomotion. Strides associated with undefined contexts (*UNDEF*), crossings performed with red traffic lights (*CWTLR*), crossings with an unidentified traffic light state (*CWTLunknown*), and events associated with smartphone use were also removed, as they were not included in the final analyses.

Subsequently, the observations were aggregated at the event level. Each event was uniquely identified through the combination of participant, locomotor context, crossing identifier, and conversation condition. For each event, the mean value of the biomechanical parameters was calculated by considering all strides belonging to that event.

Event-level aggregation was adopted for several reasons. Firstly, it allows a representative measure of the participant’s locomotor behaviour within a specific experimental condition to be obtained, reducing the influence of stride-to-stride variability. Secondly, it limits the effects due to differences in the number of available strides across events and subjects, preventing longer events from contributing disproportionately to the statistical analyses. Finally, this approach allows each observation to be directly associated with the contextual characteristics identified during the review of the egocentric videos.

The event-level aggregated dataset represented the basis for all subsequent statistical analyses.

The observation selection process led to a reduction in the overall number of crossing events available for the biomechanical analyses compared with the dataset used for the evaluation of the performance of the automatic pipeline.

In particular, the evaluation of the performance of the pedestrian crossing recognition pipeline was performed on a total of 354 manually annotated crossing events, including 91 crossings regulated by traffic lights and 263 crossings not regulated by traffic lights.

For the subsequent biomechanical analyses, crossings performed with red traffic lights (*CWTLR*), crossings with an unidentified traffic light state (*CWTLunknown*), and events associated with smartphone use were excluded. Following this selection, the final dataset included 296 crossing events, of which 67 were crossings regulated by traffic lights and 229 were crossings not regulated by traffic lights.

The final distribution of the event-level aggregated events and the corresponding number of strides across the different experimental conditions is reported in Table 3.12.

Table 3.12: Number of events and strides included in the statistical analysis, stratified by locomotor context and conversation condition.

Context	Conversation	OA events	YA events	OA strides	YA strides
Straight walking	No	37	26	5844	2367
	Yes	66	76	9279	8101
Signalised Crossing	No	14	19	689	933
	Yes	16	18	398	529
Unsignalised Crossing	No	44	35	671	491
	Yes	60	90	1085	1955

3.8 Statistical Analysis

Statistical analysis was conducted to evaluate the effects of age, locomotor context, and the presence of a conversation on the biomechanical parameters extracted during the experimental protocol. Since each participant contributed multiple observations collected in different experimental conditions, the observations could not be considered statistically independent. In addition, the number of available events varied across subjects and experimental conditions and, for some parameters, the assumption of normality was not satisfied. For these reasons, an approach based on Generalised Linear Mixed-Effects Models (GLMMs) was adopted.

3.8.1 Generalised Linear Mixed-Effects Models

For each biomechanical parameter, a Generalised Linear Mixed-Effects Model (GLMM) was estimated. This approach allows the simultaneous modelling of the effects of the experimental variables of interest and the dependence existing between observations belonging to the same subject. This aspect is particularly relevant in studies based on repeated measures, since observations from the same participant cannot be considered statistically independent.

For each biomechanical parameter, the following model was estimated:

$$Y \sim Cohort \times Context \times Conversation + (1|Subject)$$

where Y represents the parameter under analysis, while *Cohort*, *Context*, and *Conversation* represent the fixed effects included in the model.

The *Cohort* effect described the participant's cohort (*Younger Adults* or *Older Adults*), the *Context* effect represented the locomotor context, and the *Conversation* effect identified the presence or absence of a conversation during the execution of the event.

In addition to the main effects, all the possible two-way and three-way interactions between the experimental factors were included in the model. This choice made it possible to assess not only the presence of differences between cohorts or contexts, but also whether the effect of a given factor depended on the levels of the other factors.

The subject identifier was included in the model as a random effect (*random intercept*) in order to account for inter-individual variability and for the correlation between observations belonging to the same participant.

Three different locomotor contexts were considered:

- straight walking (*Straight Walking*);
- crossing regulated by a green or amber traffic lights (*Signalised Crossing*);

- crossing in the absence of traffic lights (*Unsignalised Crossing*).

The statistical distribution of the model was selected according to the characteristics of the response variable. Parameters characterised exclusively by positive values were modelled using a Gamma distribution with a logarithmic link function, which is particularly suitable for positive and asymmetric continuous variables. Parameters which did not satisfy the positivity assumption required by the Gamma distribution were instead analysed using a Normal distribution with an identity link function.

Model parameters were estimated using the Laplace approximation. The significance of the fixed effects was assessed through analysis of model deviance. The threshold for statistical significance was set at $p < 0.05$.

3.8.2 Assessment of Model Adequacy

For each estimated model, a diagnostic evaluation was performed to assess the quality of the fit to the experimental data and the absence of evident violations of the model assumptions.

For this purpose, Pearson residuals were calculated, representing the difference between observed and predicted values normalised compared with the expected variance of the model. The distribution of the residuals was examined using histograms and Quantile–Quantile plots (Q–Q plots), which were used as descriptive tools to identify possible marked deviations from the expected behaviour.

To assess the presence of possible model misspecification, residuals-versus-fitted values plots were also analysed. The absence of systematic patterns in the residuals was considered indicative of a correct description of the relationship between the explanatory variables and the response variable.

The presence of overdispersion was evaluated using the Pearson overdispersion coefficient, calculated as:

$$\phi = \frac{\sum r_i^2}{df_{res}}$$

where r_i represents the Pearson residual associated with observation i , and df_{res} indicates the residual degrees of freedom of the model. Coefficient values close to one were considered indicative of good agreement between the observed variability and that expected by the model.

To verify further the absence of unmodelled systematic structures, the correlation between Pearson residuals and fitted values was evaluated. Correlations close to zero were interpreted as evidence of the absence of residual relationships between errors and estimated values.

Finally, observed-versus-predicted plots were examined in order to assess the ability of the model to reproduce the variability observed in the experimental data and to describe correctly the general trend of the analysed parameters.

3.8.3 Post-Hoc Analysis

In the presence of statistically significant main effects or interactions, post-hoc comparisons were performed directly on the estimated GLMMs.

For each parameter, an auxiliary categorical factor was defined by combining cohort, locomotor context, and conversation condition into a single level. Comparisons between the different combinations were then performed using linear contrasts calculated directly on the coefficients estimated by the model.

In particular, the following comparisons were performed:

- between cohorts within the same combination of locomotor context and conversation condition;
- between locomotor contexts within the same cohort and the same conversation condition;
- between conversation conditions within the same cohort and the same locomotor context.

Since the high number of comparisons performed increases the risk of obtaining significant results by chance, the p -values obtained from the post-hoc comparisons were corrected using the *False Discovery Rate* (FDR) procedure, which allows control of the expected proportion of false positives among the results declared significant.

Differences were considered statistically significant for corrected p -values lower than 0.05.

All statistical analyses were performed using MATLAB R2025a (MathWorks).

Chapter 4

Results

4.1 Performance of Pre-trained Networks without and with Post-processing

4.1.1 Performance of Automatic Pedestrian Crossing Recognition

The performance of the pedestrian crossing recognition module based on Places365 was evaluated both at frame level and at event level. The frame-level evaluation was performed on all 839614 manually annotated frames, whereas the event-level evaluation was performed by comparing the automatically identified crossing episodes with those manually annotated.

Frame-Level Performance

Frame-level performance is reported in Table 4.1.

Table 4.1: Frame-level performance of raw and post-processed Places365 crosswalk detection.

Method	Accuracy	Precision	Recall	F1-score
RAW	0.871	0.870	0.593	0.705
PROCESSED	0.900	0.876	0.716	0.788

The application of temporal post-processing led to an improvement in performance compared with the raw frame-by-frame classification. In particular, accuracy increased from 0.871 to 0.900, recall from 0.593 to 0.716, and the F1-score from 0.705 to 0.788. Precision, instead, remained substantially unchanged, increasing from 0.870 for the RAW method to 0.876 for the PROCESSED method. The

application of temporal post-processing was associated with an increase in accuracy, recall, and F1-score, whereas precision remained almost unchanged.

Event-Level Performance

Event-level performance was evaluated by considering the temporal correspondence between manually annotated crossing events and automatically detected events. In particular, a bidirectional overlap threshold of 20% was used as the main criterion, computed both with respect to the duration of the manual event and to the duration of the automatic event. Since the objective of the pipeline was to identify crossing episodes to be used in the subsequent gait analysis across different locomotor contexts, the 20% overlap criterion was considered the main reference for performance evaluation. The results are reported in Table 4.2.

Table 4.2: Event-level Places365 crosswalk detection performance using a 20% bidirectional overlap criterion.

Method	Manual	Pred.	Det.	Missed	Extra	Frag.	Recall	Prec.	F1
RAW	354	6241	178	176	6022	41	0.503	0.029	0.054
PROCESSED	354	759	265	89	462	32	0.749	0.365	0.490

It should be noted that the number of predicted events and the number of detected events refer to two different counting levels. Predicted events represent the total number of crossing events automatically generated by the pipeline, whereas detected events represent the number of manual events correctly associated with at least one automatic event. Therefore, the number of predicted events does not necessarily correspond to the sum of detected and extra events, since multiple automatic fragments can be associated with the same manual event. This discrepancy is quantified by the fragmentation index, which represents the number of additional automatic fragments associated with manual events which had already been detected.

Using the 20% overlap criterion, the post-processed output correctly identified 265 of the 354 manually annotated crossing events, corresponding to a recall of 0.749. Precision increased from 0.029 for the RAW output to 0.365 for the PROCESSED output, while the F1-score increased from 0.054 to 0.490.

The number of missed events was reduced from 176 to 89, while the number of extra events decreased from 6022 to 462. In addition, the overall number of predicted events decreased from 6241 to 759. These results indicate that temporal post-processing markedly reduced the number of spurious detections and improved the ability of the pipeline to identify crossing events at the temporal level.

A sensitivity analysis was also performed by considering bidirectional overlap

thresholds equal to 10%, 20%, 30%, 50%, and 70%. The results are reported in Table 4.3.

Table 4.3: Sensitivity analysis of event-level Places365 crosswalk detection across bidirectional overlap thresholds.

Method	Overlap	Manual	Predicted	Detected	Extra	Recall	Precision	F1-score
RAW	10%	354	6241	241	5819	0.681	0.040	0.075
PROCESSED	10%	354	759	292	363	0.825	0.446	0.579
RAW	20%	354	6241	178	6022	0.503	0.029	0.054
PROCESSED	20%	354	759	265	462	0.749	0.365	0.490
RAW	30%	354	6241	112	6123	0.316	0.018	0.034
PROCESSED	30%	354	759	231	522	0.653	0.307	0.417
RAW	50%	354	6241	39	6202	0.110	0.006	0.012
PROCESSED	50%	354	759	157	602	0.444	0.207	0.282
RAW	70%	354	6241	9	6232	0.025	0.001	0.003
PROCESSED	70%	354	759	84	675	0.237	0.111	0.151

For the post-processed output, the F1-score was equal to 0.579, 0.490, 0.417, 0.282, and 0.151 for overlap thresholds of 10%, 20%, 30%, 50%, and 70%, respectively. Similarly, recall decreased from 0.825 to 0.749, 0.653, 0.444, and 0.237 as the overlap threshold increased. The highest recall and F1-score values were observed using a 10% overlap threshold, equal to 0.825 and 0.579, respectively.

As the overlap threshold increased, a progressive reduction in both recall and F1-score was observed, indicating that stricter temporal matching criteria reduced the number of events considered correctly detected.

As shown in Figure 4.1, the F1-score progressively decreased as the bidirectional overlap threshold increased.

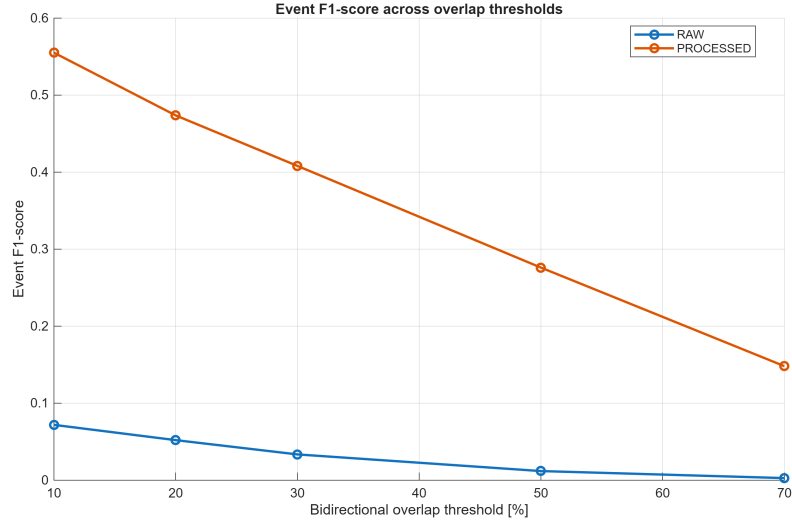


Figure 4.1: Trend of the F1-score as a function of the bidirectional overlap threshold between manual and automatically detected events.

Similarly, Figure 4.2 highlights a progressive reduction in recall as the overlap threshold increased.

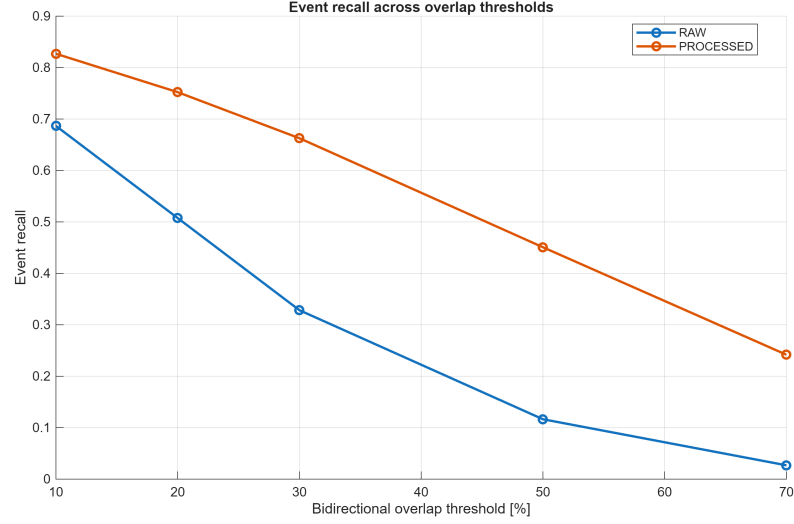


Figure 4.2: Trend of recall as a function of the bidirectional overlap threshold between manual and automatically detected events.

4.1.2 Performance of Automatic Recognition of Signalised Crossings

The performance of the YOLOv11-based module for the recognition of signalised crossings was evaluated both at frame level and at event level. The frame-level evaluation was performed on all 839 614 manually annotated frames, whereas the event-level evaluation was performed by classifying each crossing episode as either a signalised crossing or an unsignalised crossing.

Frame-Level Performance

Frame-level performance is reported in Table 4.4.

Table 4.4: Frame-level traffic-light detection performance within crosswalk frames.

Level	Total	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-score
Frame Global	839614	89668	691730	5448	52768	0.931	0.943	0.630	0.755

The frame-level classification achieved an accuracy of 0.931, a precision of 0.943, a recall of 0.630, and an F1-score of 0.755.

Event-Level Performance

Event-level performance was initially evaluated by considering a crossing event as a signalised crossing when at least 20% of the frames belonging to the event were classified as containing traffic lights. This threshold was considered the main reference for the presentation of event-level performance.

The results are reported in Table 4.5.

Table 4.5: Event-level traffic-light classification performance using 20% event threshold.

Level	Total	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-score
Event									
Crosswalk	354	81	239	24	10	0.904	0.771	0.890	0.827
Total									

Using the 20% criterion, 81 of the 91 manually annotated signalised crossing events were correctly identified, corresponding to a recall of 0.890.

Overall, 239 of the 263 unsignalised crossing events were correctly classified, whereas 24 unsignalised events were incorrectly classified as signalised crossings.

Precision was equal to 0.771 and the F1-score was equal to 0.827.

Sensitivity Analysis

A sensitivity analysis was also performed by considering classification thresholds of 10%, 20%, 30%, 40%, 50%, and 60%. The results are reported in Table 4.6.

Table 4.6: Sensitivity analysis of event-level traffic-light classification across event thresholds.

Thr	TP	TN	FP	FN	Precision	Recall	Specificity	F1-score
10%	82	228	35	9	0.701	0.901	0.867	0.788
20%	81	239	24	10	0.771	0.890	0.909	0.827
30%	76	243	20	15	0.792	0.835	0.924	0.813
40%	70	249	14	21	0.833	0.769	0.947	0.800
50%	68	254	9	23	0.883	0.747	0.966	0.810
60%	57	256	7	34	0.891	0.626	0.973	0.735

For the event-level output, recall decreased from 0.901 to 0.626 as the classification threshold increased. Similarly, the F1-score decreased from 0.788 for a 10% threshold to 0.735 for a 60% threshold.

Conversely, precision progressively increased from 0.701 to 0.891 as the classification threshold increased, while specificity increased from 0.867 to 0.973.

For the reference threshold of 20%, a recall of 0.890, a precision of 0.771 and a F1-score of 0.827 were observed.

For the events correctly classified as signalised crossings, the mean percentage of frames classified with traffic lights were 67.7%, whereas for false negative events this percentage was 2.5%. The complete descriptive values are reported in Table 4.7.

Table 4.7: Percentage of frames classified with traffic lights for correctly classified traffic-light events (TP) and false negative events (FN).

Class	N	Mean [%]	SD [%]	Median [%]	Range [%]
TP	81	67.7	18.7	71.2	25.4–97.4
FN	10	2.5	5.7	0.0	0.0–17.5

The median percentage of frames classified as including traffic lights was 71.2% for correctly classified events and 0.0% for false negative events.

As shown in Figure 4.3, recall progressively decreased as the classification threshold increased, from 0.901 for a 10% threshold to 0.626 for a 60% threshold.

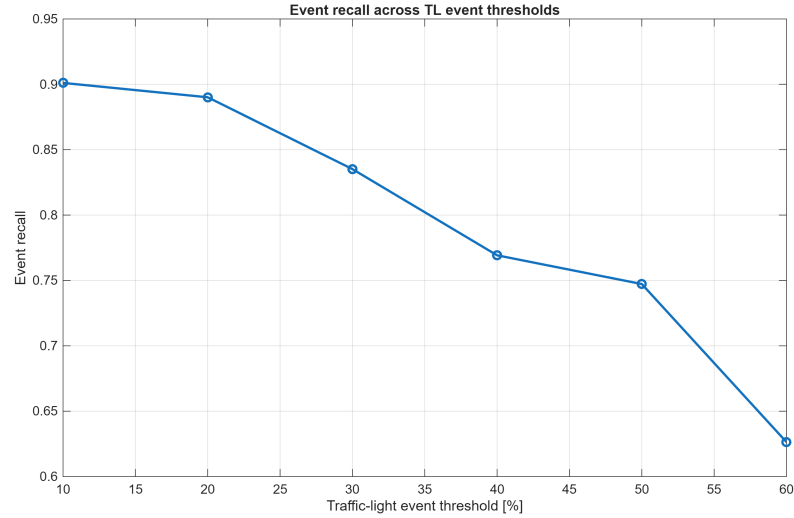


Figure 4.3: Trend of recall as a function of the threshold used to classify an event as a signalised crossing.

Similarly, Figure 4.4 shows an overall reduction in the F1-score, from 0.788 for a 10% threshold to 0.735 for a 60% threshold.

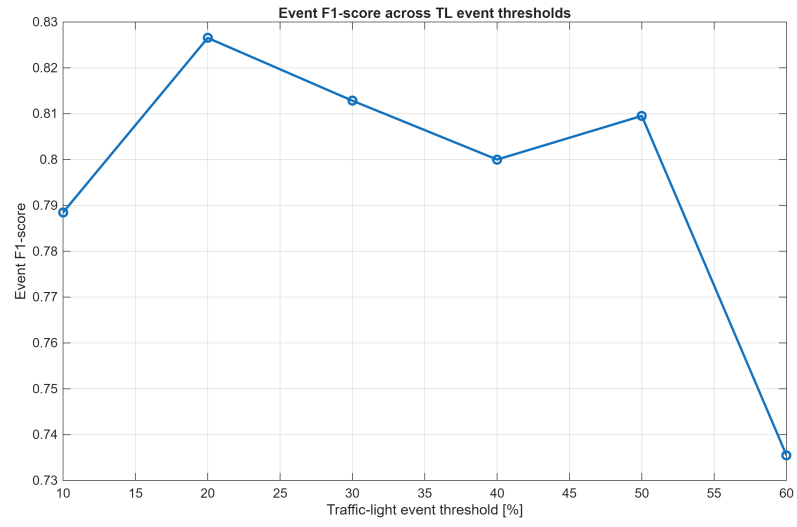


Figure 4.4: Trend of the F1-score as a function of the threshold used to classify an event as a signalised crossing.

4.2 Aggregated Analysis

As far as a first analysis is concerned, all the events were aggregated independently of the locomotor context and the presence of conversation, in order to evaluate the possible differences between the two cohorts without considering the influence of environmental conditions. For each parameter, a Generalised Linear Mixed-Effects Model (GLMM) was therefore estimated, including only the cohort as a fixed effect and the subject as a random effect:

$$Y \sim \text{Cohort} + (1|\text{Subject})$$

The results of the aggregated analysis are reported in Table 4.8.

Table 4.8: Aggregated GLMM results.

Parameter	Distribution	Cohort p-value	Result
Froude number	Gamma	0.081	Not significant
stride speed	Gamma	0.078	Not significant
stride length	Gamma	0.552	Not significant
stride duration	Gamma	0.322	Not significant
cadence	Gamma	0.298	Not significant
aRMS/speed V-HD	Gamma	0.852	Not significant
aRMS/speed ML-HD	Gamma	0.726	Not significant
aRMS/speed AP-HD	Gamma	0.878	Not significant
RCME V-HD	Gamma	0.001	Significant
RCME ML-HD	Gamma	0.069	Not significant
RCME AP-HD	Gamma	0.143	Not significant
LDLJ V-HD	Normal	0.131	Not significant
LDLJ ML-HD	Normal	0.543	Not significant
LDLJ AP-HD	Normal	0.240	Not significant

The aggregated analysis showed a significant difference between the two cohorts only for RCME V-HD ($p = 0.001$). None of the spatiotemporal gait parameters showed a significant cohort effect. Similarly, no significant differences emerged for the normalised magnitude parameters of head movements, for the other head entropy parameters, namely RCME ML-HD and RCME AP-HD, or for the smoothness parameters.

The aggregated distribution of RCME V-HD in the two cohorts is reported in Figure 4.5. Overall, the aggregated analysis showed a limited ability to highlight the differences between younger adults and older adults, identifying a significant cohort effect only for the vertical regularity of head movements.

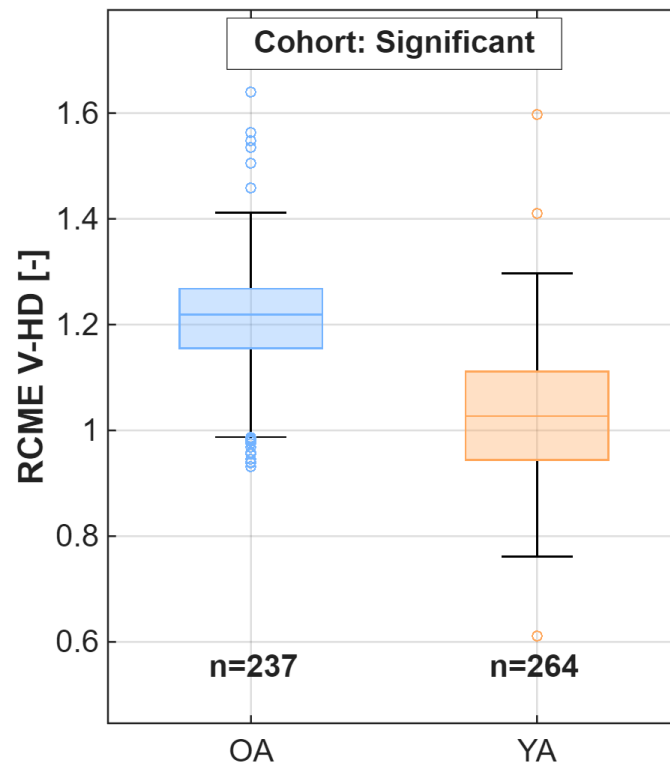


Figure 4.5: Aggregated distribution of RCME V-HD, the only parameter found to be significantly different between the two cohorts in the aggregated analysis.

4.3 Context-Aware Analysis

Subsequently, a context-aware analysis was performed, keeping events distinct according to the locomotor context and the presence or the absence of face-to-face conversation. The contexts which were considered were straight walking, signalised crossing, and unsignalised crossing.

For each parameter, a Generalised Linear Mixed-Effects Model (GLMM) was estimated, including cohort (*Cohort*), locomotor context (*Context*), and conversation (*Conversation*) as fixed effects, together with all related interactions. The subject was instead included as a random effect:

$$Y \sim \text{Cohort} \times \text{Context} \times \text{Conversation} + (1|\text{Subject})$$

The results are presented separately for the spatiotemporal gait parameters, the magnitude of head movements, the regularity of head movements, and the smoothness of head movements. For each parameter, descriptive values are reported as median and interquartile range (IQR) of the means calculated at the event level, together with the results of the statistical model. For some parameters, graphical representations of the distributions across the different experimental conditions are also reported.

4.3.1 Spatiotemporal Gait Parameters

Spatiotemporal gait parameters were analysed in order to evaluate the effect of the cohort, the locomotor context, and the conversation on the characteristics of real-world locomotion. The descriptive and inferential results obtained using GLMMs are reported in the tables dedicated to each parameter.

Froude Number

Froude number values across the different locomotor contexts and conversation conditions are reported in Table 4.9 and Figure 4.6.

Table 4.9: Froude number [-]. Descriptive values are reported as median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

Froude number [-]			
Context	Conversation	OA	YA
Straight walking	No	0.226 (0.031)	0.196 (0.035)
	Yes	0.235 (0.043)	0.200 (0.037)
Signalised crossing	No	0.254 (0.039)	0.238 (0.040)
	Yes	0.217 (0.063)	0.210 (0.050)
Unsignalised crossing	No	0.214 (0.048)	0.207 (0.058)
	Yes	0.233 (0.046)	0.205 (0.044)
GLMM results:			
Cohort		0.031	
Context		<0.001	
Conversation		0.665	
Cohort $\times$ Context		0.133	
Cohort $\times$ Conversation		0.983	
Context $\times$ Conversation		0.002	
Cohort $\times$ Context $\times$ Conversation		0.714	

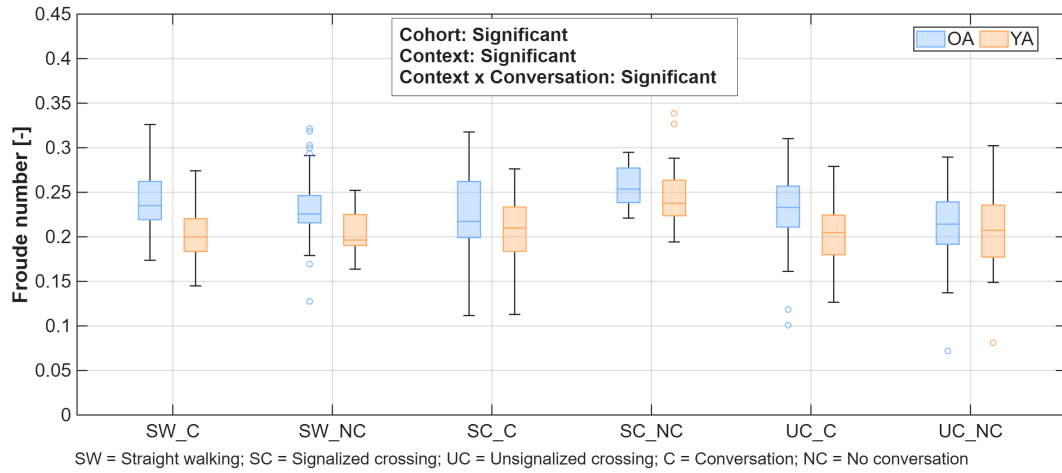


Figure 4.6: Distribution of Froude number values across the different locomotor contexts and conversation conditions, considered separately for each cohort.

The GLMM analysis showed a significant effect of cohort ($p = 0.031$) and locomotor context ($p < 0.001$), as well as a significant Context $\times$ Conversation interaction ($p = 0.002$). No other significant interactions emerged.

The FDR-corrected post-hoc analyses (Table 4.23) showed that, in the absence of conversation, the Froude number was significantly higher during signalised crossings

than during unsignalised crossings in both cohorts. In older adults, significant differences were also observed between straight walking and unsignalised crossings both in the presence and in the absence of conversation. Finally, during signalised crossings, the Froude number was significantly higher in the absence of conversation than in the conversation condition in both cohorts.

Stride Speed

Stride speed values in the different locomotor contexts and conversation conditions are reported in Table 4.10 and Figure 4.7.

Table 4.10: Stride speed [m/s]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

Stride speed [m/s]			
Context	Conversation	OA	YA
Straight walking	No	1.409 (0.089)	1.327 (0.096)
	Yes	1.460 (0.142)	1.329 (0.109)
Signalised crossing	No	1.530 (0.091)	1.459 (0.111)
	Yes	1.380 (0.215)	1.358 (0.161)
Unsignalised crossing	No	1.372 (0.174)	1.351 (0.168)
	Yes	1.457 (0.165)	1.345 (0.144)
GLMM results:			
Cohort		0.034	
Context		<0.001	
Conversation		0.770	
Cohort $\times$ Context		0.175	
Cohort $\times$ Conversation		0.925	
Context $\times$ Conversation		<0.001	
Cohort $\times$ Context $\times$ Conversation		0.649	

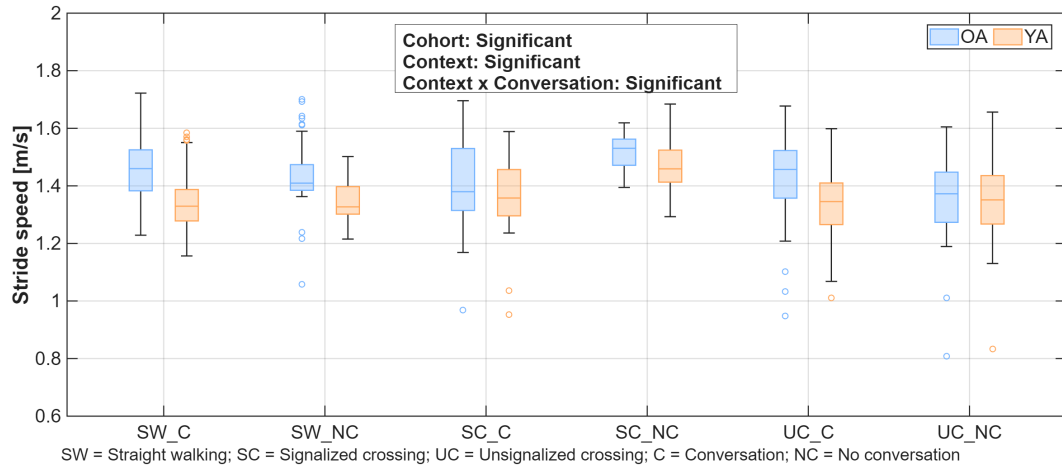


Figure 4.7: Distribution of stride speed values in the different locomotor contexts and conversation conditions, considered separately for each cohort.

The GLMM analysis showed a significant effect of cohort ($p = 0.034$) and locomotor context ($p < 0.001$), as well as a significant Context $\times$ Conversation interaction ($p < 0.001$). In contrast, no significant effect of conversation as a main effect nor significant Cohort $\times$ Context, Cohort $\times$ Conversation, or Cohort $\times$ Context $\times$ Conversation interactions, were observed.

The FDR-corrected post-hoc analyses (Table 4.23) showed that, in older adults in the absence of conversation, stride speed was significantly higher during signalised crossings than during unsignalised crossings. In the same cohort, significant differences were also observed between straight walking and unsignalised crossings both in the absence and in the presence of conversation.

In younger adults, in the absence of conversation, stride speed was significantly higher during signalised crossings than during both straight walking and unsignalised crossings. Finally, during signalised crossings, stride speed was significantly higher in the absence of conversation than in the conversation condition in both cohorts.

Stride Length

Stride length values in the different locomotor contexts and conversation conditions are reported in Table 4.11 and Figure 4.8.

Table 4.11: Stride length [m]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

Stride length [m]			
Context	Conversation	OA	YA
Straight walking	No	1.446 (0.165)	1.463 (0.100)
	Yes	1.529 (0.141)	1.481 (0.205)
Signalised crossing	No	1.566 (0.194)	1.579 (0.195)
	Yes	1.466 (0.245)	1.498 (0.186)
Unsignalised crossing	No	1.440 (0.213)	1.458 (0.158)
	Yes	1.529 (0.144)	1.495 (0.167)
GLMM results:			
Cohort		0.301	
Context		<0.001	
Conversation		0.686	
Cohort $\times$ Context		0.048	
Cohort $\times$ Conversation		0.856	
Context $\times$ Conversation		0.002	
Cohort $\times$ Context $\times$ Conversation		0.462	

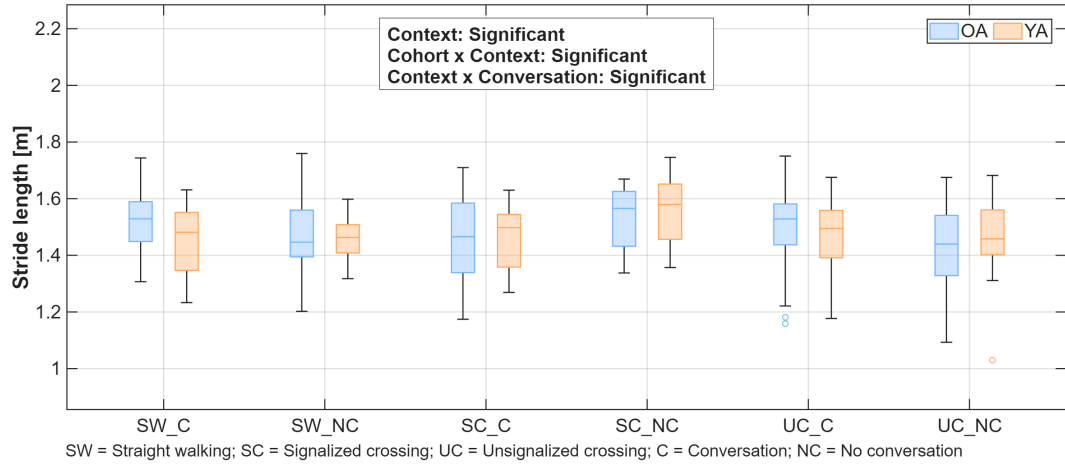


Figure 4.8: Distribution of stride length values in the different locomotor contexts and conversation conditions, considered separately for each cohort.

The GLMM analysis showed a significant effect of locomotor context ($p < 0.001$), as well as significant Cohort $\times$ Context ($p = 0.048$) and Context $\times$ Conversation ($p = 0.002$) interactions. In contrast, no significant main effects of cohort ($p = 0.301$) or conversation ($p = 0.686$), nor other significant interactions emerged.

The FDR-corrected post-hoc analyses (Table 4.23) showed that, in older adults in the absence of conversation, stride length was significantly greater during straight walking and signalised crossings than during unsignalised crossings. In older adults in the presence of conversation, significant differences were observed between straight walking and signalised crossings, and between straight walking and unsignalised crossings.

In younger adults, in the absence of conversation, stride length was significantly greater during signalised crossings than during both straight walking and unsignalised crossings. Finally, during signalised crossings, stride length was significantly greater in the absence of conversation than in the conversation condition in both cohorts.

Stride Duration

Stride duration values in the different locomotor contexts and conversation conditions are reported in Table 4.12 and Figure 4.9.

Table 4.12: Stride duration [s]. Descriptive values are reported as median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

Context	Stride duration [s]		
	Conversation	OA	YA
Straight walking	No	1.026 (0.062)	1.086 (0.173)
	Yes	1.052 (0.045)	1.059 (0.100)
Signalised crossing	No	1.025 (0.045)	1.050 (0.150)
	Yes	1.063 (0.049)	1.056 (0.135)
Unsignalised crossing	No	1.043 (0.080)	1.074 (0.193)
	Yes	1.057 (0.055)	1.060 (0.141)
GLMM results:			
Cohort		0.308	
Context		<0.001	
Conversation		0.928	
Cohort × Context		0.801	
Cohort × Conversation		0.977	
Context × Conversation		0.001	
Cohort × Context × Conversation		0.864	

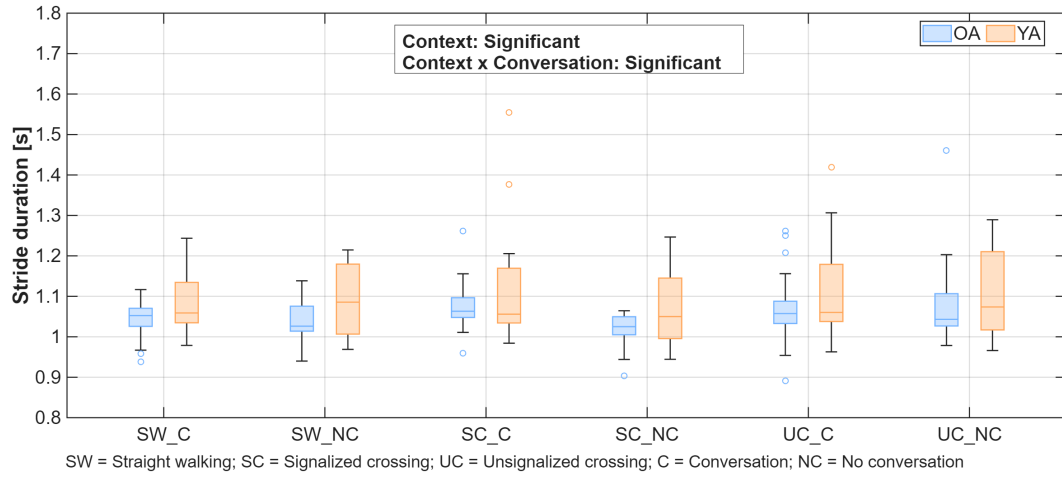


Figure 4.9: Distribution of stride duration values in the different locomotor contexts and conversation conditions, considered separately for each cohort.

The GLMM analysis showed a significant main effect of locomotor context ($p < 0.001$) and a significant Context $\times$ Conversation interaction ($p = 0.001$). In contrast, no significant main effects of cohort ($p = 0.308$) or conversation ($p = 0.928$) were found. Similarly, no significant Cohort $\times$ Context, Cohort $\times$ Conversation, or Cohort $\times$ Context $\times$ Conversation interactions emerged.

The FDR-corrected post-hoc analyses (Table 4.23) showed that, in older adults in the absence of conversation, stride duration was significantly longer during unsignalised crossings than during both straight walking and signalised crossings. In older adults in the presence of conversation, stride duration was significantly longer during signalised crossings than during straight walking.

In younger adults, in the absence of conversation, stride duration was significantly longer during unsignalised crossings than during signalised crossings. Finally, during signalised crossings, stride duration was significantly longer in the presence of conversation than in the no-conversation condition in both cohorts.

Cadence

Cadence values across the different locomotor contexts and conversation conditions are reported in Table 4.13 and Figure 4.10.

Table 4.13: Cadence [steps/min]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

Context	Cadence [steps/min]		
	Conversation	OA	YA
Straight walking	No	117.046 (6.850)	110.697 (17.461)
	Yes	114.097 (4.861)	113.513 (10.278)
Signalised crossing	No	117.284 (5.097)	114.503 (15.094)
	Yes	113.211 (4.486)	113.901 (13.523)
Unsignalised crossing	No	115.622 (8.998)	111.989 (18.963)
	Yes	113.759 (5.897)	113.360 (13.514)
GLMM results:			
Cohort		0.279	
Context		<0.001	
Conversation		0.876	
Cohort $\times$ Context		0.911	
Cohort $\times$ Conversation		0.957	
Context $\times$ Conversation		0.003	
Cohort $\times$ Context $\times$ Conversation		0.972	

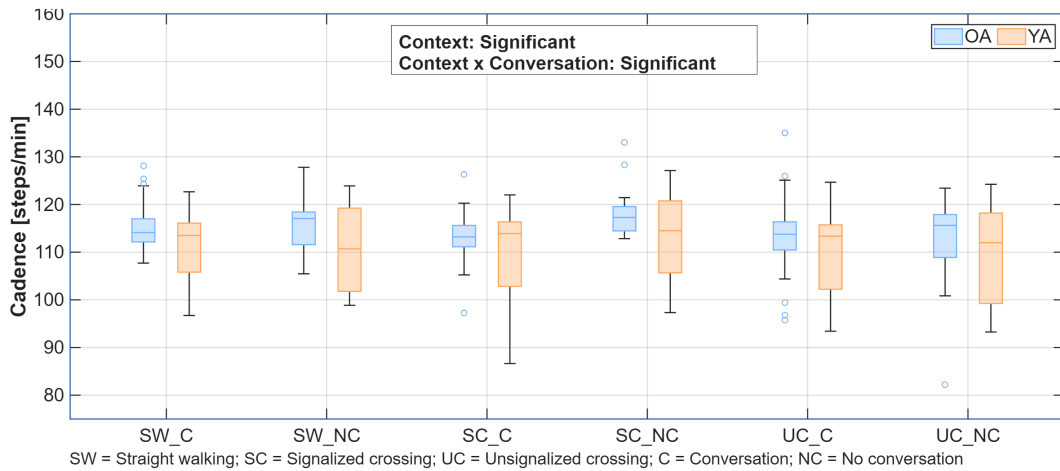


Figure 4.10: Distribution of cadence values in the different locomotor contexts and conversation conditions, considered separately for each cohort.

The GLMM analysis showed a significant effect of locomotor context ($p < 0.001$) and a significant Context $\times$ Conversation interaction ($p = 0.003$). In contrast, no significant effects of cohort, conversation as a main effect, or the remaining interactions were observed.

The FDR-corrected post-hoc analyses (Table 4.23) showed that, in older adults in the absence of conversation, cadence was significantly higher during straight walking and signalised crossings than during unsignalised crossings. In younger adults, in the absence of conversation, cadence was significantly higher during signalised crossings than during both straight walking and unsignalised crossings. Finally, during signalised crossings, cadence was significantly higher in the absence of conversation than in the conversation condition in both cohorts.

Summary of Spatiotemporal Gait Parameters. Overall, all spatiotemporal gait parameters showed a significant effect of locomotor context and a significant Context $\times$ Conversation interaction. In addition, the Froude number and stride speed showed a significant main effect of cohort. Stride length also showed a significant Cohort $\times$ Context interaction. No significant Cohort $\times$ Conversation interactions or three-way interactions emerged.

4.3.2 Head-Movement Parameters

In addition to spatiotemporal gait parameters, several descriptive parameters of head movements were analysed in order to evaluate possible differences associated with cohort, locomotor context, and the presence of conversation. In particular, measures related to the magnitude, regularity, and smoothness of head movements were considered. The obtained results are reported in the following sections.

Magnitude

The magnitude of head movements was quantified using the RMS value of head acceleration normalised compared to walking speed (aRMS/speed). The descriptive and inferential results are reported in Table 4.14, Table 4.15, Table 4.16, and Figure 4.11.

Table 4.14: aRMS/speed V-HD [s^{-1}]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

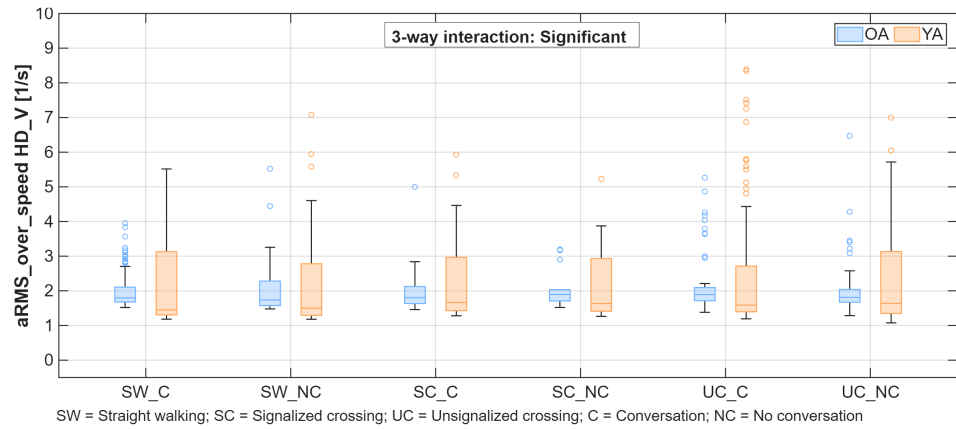
aRMS/speed V-HD [s^{-1}]			
Context	Conversation	OA	YA
Straight walking	No	1.739 (0.706)	1.499 (1.490)
	Yes	1.797 (0.428)	1.449 (1.824)
Signalised crossing	No	1.897 (0.329)	1.633 (1.519)
	Yes	1.805 (0.491)	1.665 (1.537)
Unsignalised crossing	No	1.811 (0.369)	1.639 (1.785)
	Yes	1.894 (0.381)	1.589 (1.319)
GLMM results:			
Cohort		0.746	
Context		0.749	
Conversation		0.760	
Cohort $\times$ Context		0.378	
Cohort $\times$ Conversation		0.161	
Context $\times$ Conversation		0.922	
Cohort $\times$ Context $\times$ Conversation		0.041	

Table 4.15: aRMS/speed ML-HD [s^{-1}]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

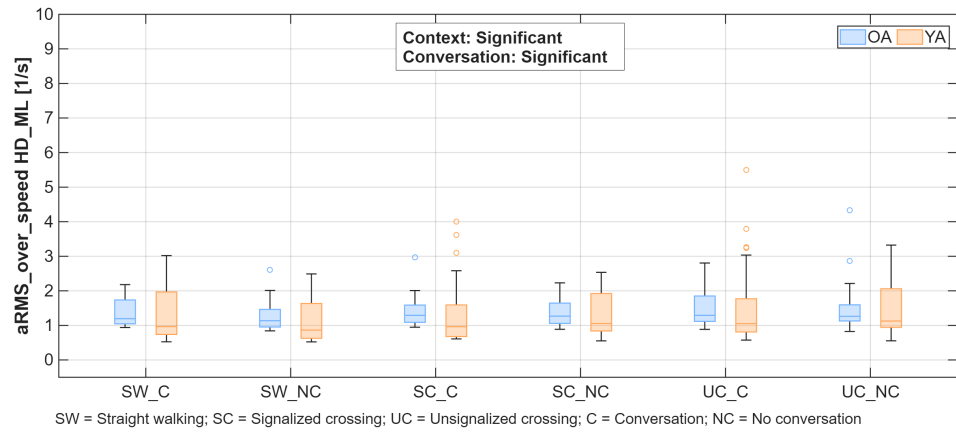
aRMS/speed ML-HD [s^{-1}]			
Context	Conversation	OA	YA
Straight walking	No	1.135 (0.510)	0.865 (1.007)
	Yes	1.194 (0.689)	0.967 (1.229)
Signalised crossing	No	1.267 (0.587)	1.052 (1.083)
	Yes	1.289 (0.499)	0.963 (0.915)
Unsignalised crossing	No	1.264 (0.467)	1.126 (1.117)
	Yes	1.290 (0.734)	1.051 (0.960)
GLMM results:			
Cohort		0.735	
Context		0.003	
Conversation		0.004	
Cohort $\times$ Context		0.790	
Cohort $\times$ Conversation		0.491	
Context $\times$ Conversation		0.377	
Cohort $\times$ Context $\times$ Conversation		0.811	

Table 4.16: aRMS/speed AP-HD [s^{-1}]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

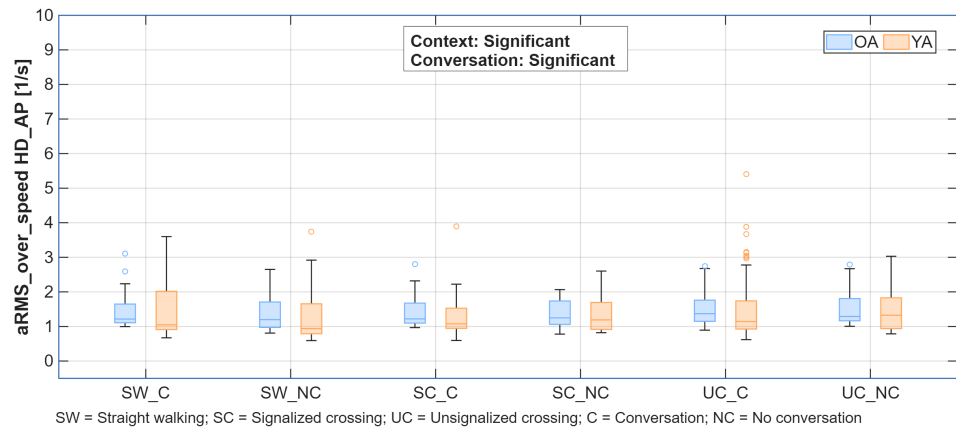
aRMS/speed AP-HD [s^{-1}]			
Context	Conversation	OA	YA
Straight walking	No	1.198 (0.736)	0.937 (0.867)
	Yes	1.214 (0.542)	1.047 (1.108)
Signalised crossing	No	1.249 (0.677)	1.189 (0.778)
	Yes	1.219 (0.579)	1.076 (0.584)
Unsignalised crossing	No	1.289 (0.645)	1.325 (0.891)
	Yes	1.370 (0.612)	1.142 (0.814)
GLMM results:			
Cohort		0.872	
Context		0.010	
Conversation		0.048	
Cohort $\times$ Context		0.892	
Cohort $\times$ Conversation		0.952	
Context $\times$ Conversation		0.469	
Cohort $\times$ Context $\times$ Conversation		0.930	



(a) aRMS/speed V-HD.



(b) aRMS/speed ML-HD.



(c) aRMS/speed AP-HD.

Figure 4.11: Distribution of aRMS/speed values in the different locomotor contexts and conversation conditions, considered separately for each cohort: (a) vertical component, (b) medio-lateral component, and (c) anterior-posterior component.

For the vertical component of head acceleration normalised compared to walking speed (aRMS/speed V-HD), the GLMM analysis did not show significant main effects of cohort ($p = 0.746$), locomotor context ($p = 0.749$), or conversation ($p = 0.760$). No significant Cohort $\times$ Context, Cohort $\times$ Conversation, or Context $\times$ Conversation interactions were observed. However, a significant three-way Cohort $\times$ Context $\times$ Conversation interaction was observed ($p = 0.041$). The FDR-corrected post-hoc analyses (Table 4.23) showed, in younger adults in the presence of conversation, significant differences between straight walking and signalised crossings, and between straight walking and unsignalised crossings. In younger adults, a significant difference between conditions with and without conversation during unsignalised crossings was also observed.

For the medio-lateral component (aRMS/speed ML-HD), the GLMM analysis showed a significant effect of locomotor context ($p = 0.003$) and conversation ($p = 0.004$). No significant cohort effect ($p = 0.735$) or significant interactions between the considered factors were observed. The FDR-corrected post-hoc comparisons showed significant differences between straight walking and unsignalised crossings in both cohorts in the absence of conversation. In younger adults in the presence of conversation, significant differences were also observed between straight walking and signalised crossings, and between straight walking and unsignalised crossings. In older adults, a significant difference between conditions with and without conversation during straight walking was observed.

For the anterior-posterior component (aRMS/speed AP-HD), a significant effect of locomotor context ($p = 0.010$) and a significant effect of conversation ($p = 0.048$) emerged. No significant cohort effect ($p = 0.872$), nor significant Cohort $\times$ Context, Cohort $\times$ Conversation, Context $\times$ Conversation, or Cohort $\times$ Context $\times$ Conversation interactions were observed. The FDR-corrected post-hoc analyses showed a significant difference between straight walking and unsignalised crossings in older adults in the absence of conversation.

Overall, the results related to the magnitude of head movements showed significant effects of locomotor context and conversation for the medio-lateral and anterior-posterior components, whereas for the vertical component only a significant Cohort $\times$ Context $\times$ Conversation interaction emerged.

Regularity

The regularity of head movements was evaluated using Refined Composite Multiscale Entropy (RCME). The descriptive and inferential results are reported in Table 4.17, Table 4.18, Table 4.19, and Figure 4.12.

Table 4.17: RCME V-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

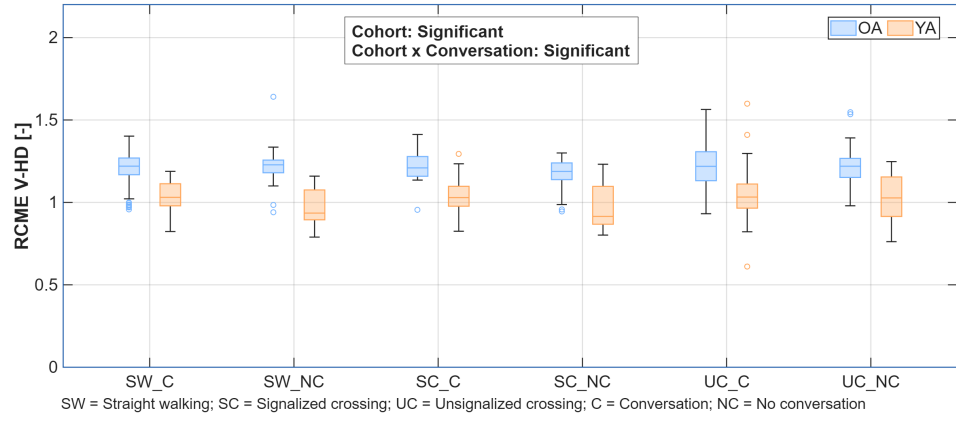
RCME V-HD [-]			
Context	Conversation	OA	YA
Straight walking	No	1.228 (0.077)	0.935 (0.181)
	Yes	1.220 (0.101)	1.031 (0.134)
Signalised crossing	No	1.188 (0.102)	0.915 (0.230)
	Yes	1.209 (0.119)	1.029 (0.121)
Unsignalised crossing	No	1.220 (0.116)	1.027 (0.240)
	Yes	1.219 (0.176)	1.032 (0.146)
GLMM results:			
Cohort		<0.001	
Context		0.528	
Conversation		0.877	
Cohort × Context		0.130	
Cohort × Conversation		0.013	
Context × Conversation		0.903	
Cohort × Context × Conversation		0.108	

Table 4.18: RCME ML-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

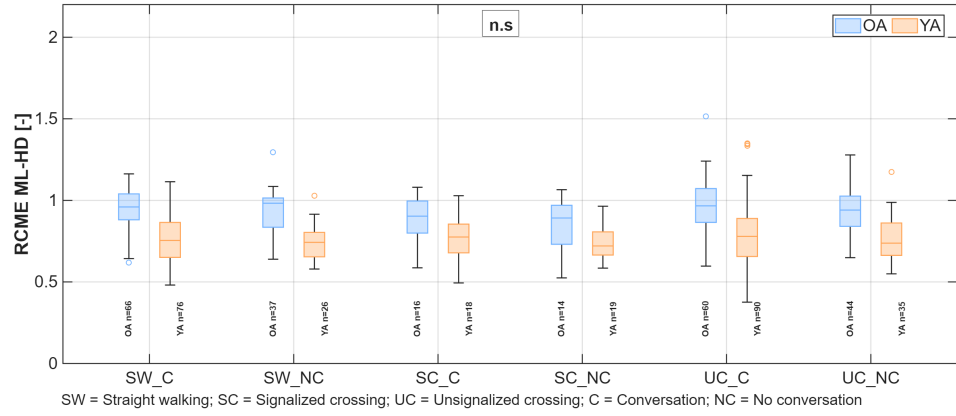
RCME ML-HD [-]			
Context	Conversation	OA	YA
Straight walking	No	0.983 (0.179)	0.743 (0.150)
	Yes	0.959 (0.159)	0.754 (0.214)
Signalised crossing	No	0.892 (0.239)	0.720 (0.142)
	Yes	0.903 (0.197)	0.775 (0.176)
Unsignalised crossing	No	0.940 (0.186)	0.738 (0.199)
	Yes	0.966 (0.208)	0.779 (0.233)
GLMM results:			
Cohort		0.055	
Context		0.059	
Conversation		0.537	
Cohort × Context		0.172	
Cohort × Conversation		0.853	
Context × Conversation		0.522	
Cohort × Context × Conversation		0.328	

Table 4.19: Head RCME AP-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

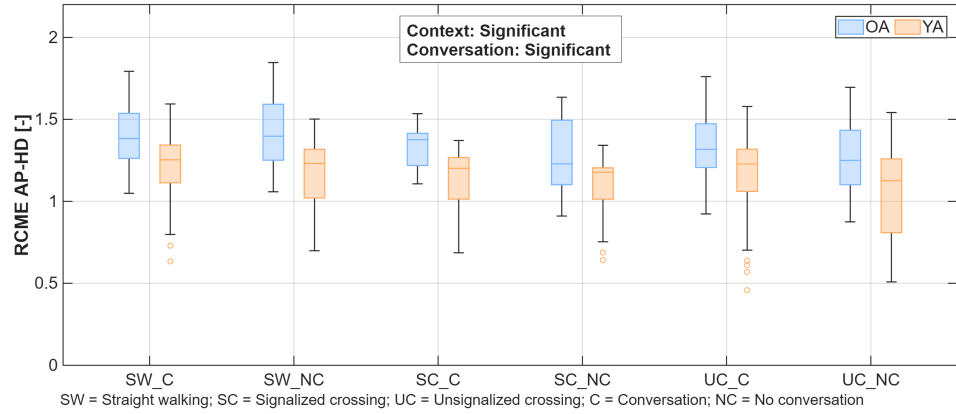
RCME AP-HD [-]			
Context	Conversation	OA	YA
Straight walking	No	1.397 (0.341)	1.231 (0.298)
	Yes	1.383 (0.275)	1.253 (0.230)
Signalised crossing	No	1.229 (0.393)	1.178 (0.191)
	Yes	1.376 (0.196)	1.202 (0.254)
Unsignalised crossing	No	1.250 (0.333)	1.126 (0.450)
	Yes	1.317 (0.267)	1.227 (0.258)
GLMM results:			
Cohort		0.242	
Context		< 0.001	
Conversation		0.029	
Cohort × Context		0.793	
Cohort × Conversation		0.193	
Context × Conversation		0.182	
Cohort × Context × Conversation		0.869	



(a) RCME V-HD.



(b) RCME ML-HD.



(c) RCME AP-HD.

Figure 4.12: Distribution of RCME values across the different locomotor contexts and conversation conditions, considered separately for each cohort: (a) vertical component, (b) medio-lateral component, and (c) anterior-posterior component.

For RCME V-HD, a significant main effect of cohort emerged ($p < 0.001$), together with a significant Cohort $\times$ Conversation interaction ($p = 0.013$). No significant effects of locomotor context ($p = 0.528$), conversation as a main effect ($p = 0.877$), or Cohort $\times$ Context, Context $\times$ Conversation, or Cohort $\times$ Context $\times$ Conversation interactions were observed. The FDR-corrected post-hoc analyses (Table 4.23) showed significant differences between the two cohorts in all combinations of locomotor context and conversation. In younger adults, significant differences were also observed between straight walking and unsignalised crossings in the absence of conversation, and between conditions with and without conversation during straight walking.

For RCME ML-HD, no significant effects of cohort ($p = 0.055$), locomotor context ($p = 0.059$), or conversation ($p = 0.537$) emerged. No significant interactions between the considered factors were observed. However, the FDR-corrected post-hoc analyses showed, in older adults in the presence of conversation, significant differences between straight walking and signalised crossings, and between signalised and unsignalised crossings.

For RCME AP-HD, a significant effect of locomotor context ($p < 0.001$) and a significant effect of conversation ($p = 0.029$) emerged. No significant effects of cohort ($p = 0.242$), nor significant Cohort $\times$ Context, Cohort $\times$ Conversation, Context $\times$ Conversation, or Cohort $\times$ Context $\times$ Conversation interactions were observed. The FDR-corrected post-hoc analyses (Table 4.23) showed significant differences between specific locomotor contexts in both older adults and younger adults. In particular, in older adults, differences were observed between straight walking and unsignalised crossings in the absence of conversation, and between straight walking and both signalised and unsignalised crossings in the presence of conversation. In younger adults, differences were observed between straight walking and the other locomotor contexts both in the absence and in the presence of conversation. In older adults, a significant difference between conditions with and without conversation during unsignalised crossings was also observed.

Overall, for RCME V-HD, a significant cohort effect was observed, together with a significant Cohort $\times$ Conversation interaction. For RCME AP-HD, significant effects of locomotor context and conversation emerged. For RCME ML-HD, no significant effects were observed in the GLMM model, although some post-hoc comparisons were significant in older adults in the presence of conversation.

Smoothness

The smoothness of head movements was quantified using the Log Dimensionless Jerk (LDLJ). The descriptive and inferential results are reported in Table 4.20, Table 4.21, Table 4.22, and Figure 4.13.

Table 4.20: LDLJ V-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

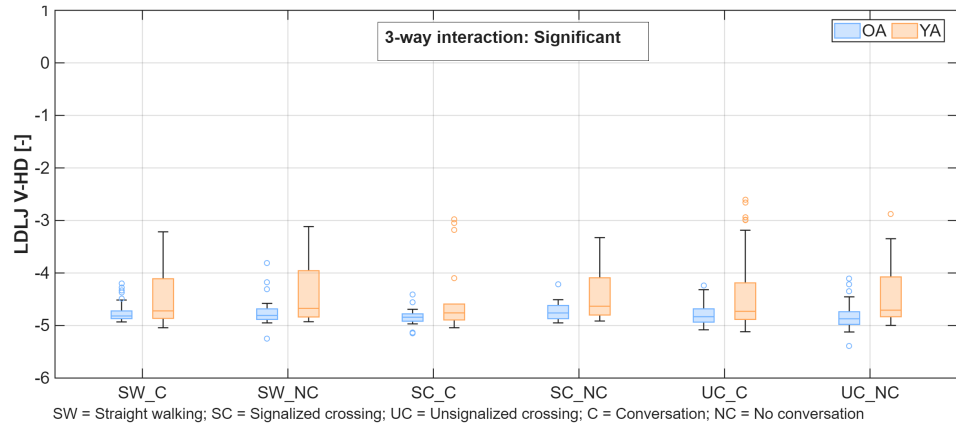
LDLJ V-HD [-]			
Context	Conversation	OA	YA
Straight walking	No	-4.809 (0.201)	-4.674 (0.881)
	Yes	-4.818 (0.146)	-4.722 (0.754)
Signalised crossing	No	-4.761 (0.251)	-4.633 (0.709)
	Yes	-4.842 (0.140)	-4.759 (0.303)
Unsignalised crossing	No	-4.872 (0.242)	-4.708 (0.758)
	Yes	-4.832 (0.253)	-4.730 (0.696)
GLMM results:			
Cohort		0.101	
Context		0.293	
Conversation		0.715	
Cohort×Context		0.329	
Cohort×Conversation		0.155	
Context×Conversation		0.681	
Cohort×Context×Conversation		0.028	

Table 4.21: LDLJ ML-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

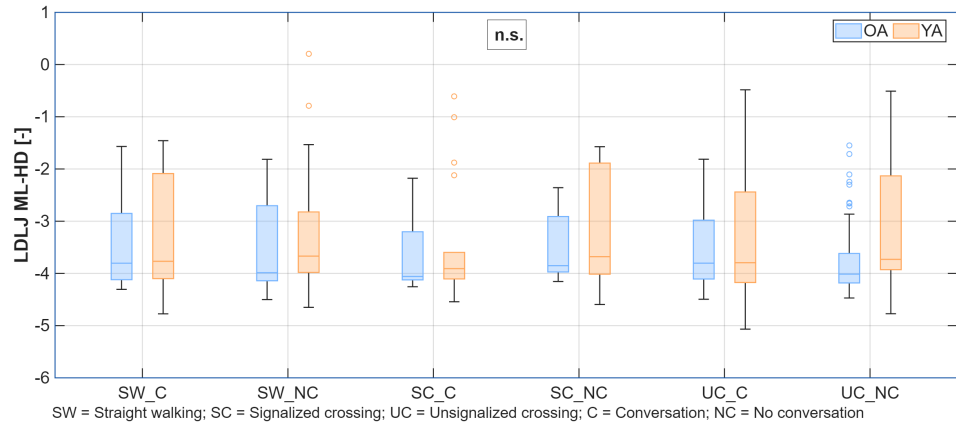
LDLJ ML-HD [-]			
Context	Conversation	OA	YA
Straight walking	No	-3.987 (1.437)	-3.668 (1.158)
	Yes	-3.804 (1.269)	-3.768 (2.012)
Signalised crossing	No	-3.852 (1.062)	-3.680 (2.127)
	Yes	-4.058 (0.922)	-3.909 (0.509)
Unsignalised crossing	No	-4.013 (0.564)	-3.730 (1.797)
	Yes	-3.804 (1.127)	-3.795 (1.734)
GLMM results:			
Cohort		0.424	
Context		0.364	
Conversation		0.889	
Cohort×Context		0.431	
Cohort×Conversation		0.111	
Context×Conversation		0.531	
Cohort×Context×Conversation		0.229	

Table 4.22: LDLJ AP-HD [-]. Descriptive values are reported as the median (IQR) of event-level mean values. GLMM p-values are reported below the descriptive rows.

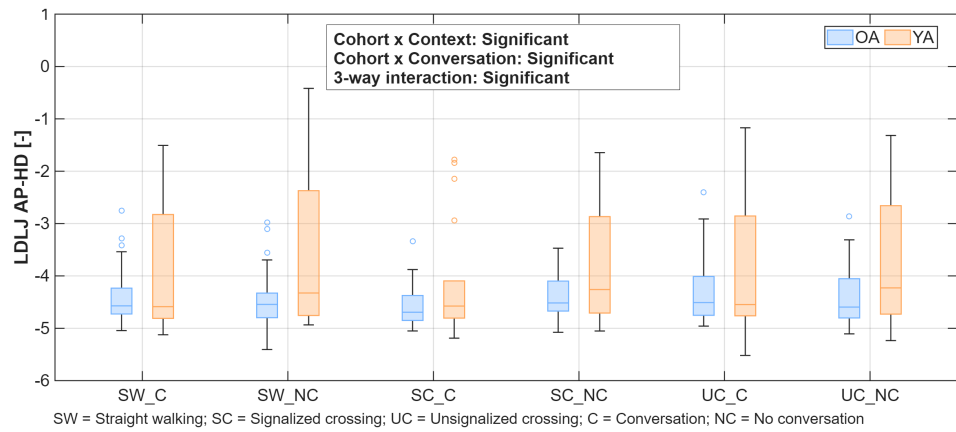
LDLJ AP-HD [-]			
Context	Conversation	OA	YA
Straight walking	No	-4.543 (0.469)	-4.326 (2.383)
	Yes	-4.571 (0.494)	-4.587 (1.983)
Signalised crossing	No	-4.516 (0.573)	-4.260 (1.842)
	Yes	-4.694 (0.477)	-4.576 (0.710)
Unsignalised crossing	No	-4.595 (0.749)	-4.229 (2.071)
	Yes	-4.510 (0.741)	-4.547 (1.907)
GLMM results:			
Cohort		0.069	
Context		0.249	
Conversation		0.126	
Cohort $\times$ Context		<0.001	
Cohort $\times$ Conversation		<0.001	
Context $\times$ Conversation		0.177	
Cohort $\times$ Context $\times$ Conversation		<0.001	



(a) LDLJ V-HD.



(b) LDLJ ML-HD.



(c) LDLJ AP-HD.

Figure 4.13: Distribution of LDLJ values in the different locomotor contexts and conversation conditions, considered separately for each cohort: (a) vertical component, (b) medio-lateral component, and (c) anterior-posterior component.

For LDLJ V-HD, the GLMM analysis did not show significant main effects of cohort ($p = 0.101$), locomotor context ($p = 0.293$), or conversation ($p = 0.715$). No significant Cohort $\times$ Context, Cohort $\times$ Conversation, or Context $\times$ Conversation interactions were observed. However, a significant three-way Cohort $\times$ Context $\times$ Conversation interaction was observed ($p = 0.028$). The FDR-corrected post-hoc analyses (Table 4.23) showed, in younger adults, a significant difference between straight walking and unsignalised crossings in the absence of conversation, and a significant difference between conditions with and without conversation during unsignalised crossings.

For LDLJ ML-HD, no significant effects of cohort ($p = 0.424$), locomotor context ($p = 0.364$), or conversation ($p = 0.889$) were observed. No significant interactions between the considered factors emerged.

For LDLJ AP-HD, no significant main effects of cohort ($p = 0.069$), locomotor context ($p = 0.249$), or conversation ($p = 0.126$) emerged. Significant Cohort $\times$ Context ($p < 0.001$), Cohort $\times$ Conversation ($p < 0.001$), and Cohort $\times$ Context $\times$ Conversation ($p < 0.001$) interactions were instead observed. The Context $\times$ Conversation interaction was not significant ($p = 0.177$). The FDR-corrected post-hoc analyses showed, in younger adults in the absence of conversation, significant differences between straight walking and signalised crossings and between straight walking and unsignalised crossings. In younger adults, a significant difference between conditions with and without conversation during straight walking was also observed.

Overall, for LDLJ V-HD, only a significant Cohort $\times$ Context $\times$ Conversation interaction emerged. For LDLJ ML-HD, no significant effects were observed. For LDLJ AP-HD, significant interactions involving cohort, locomotor context, and conversation were observed, in the absence of significant main effects.

Summary of Head-Movement Parameters. The context-aware analysis of head-movement parameters highlighted different effects depending on the metric considered and the direction analysed. Regarding the magnitude of head movements, the vertical component aRMS/speed V-HD showed only a significant Cohort $\times$ Context $\times$ Conversation interaction. The medio-lateral and anterior-posterior components instead showed significant effects of locomotor context and conversation, in the absence of significant cohort effects or significant interactions.

Regarding the regularity of head movements, RCME V-HD showed a significant main effect of cohort and a significant Cohort $\times$ Conversation interaction. RCME ML-HD did not show significant effects in the GLMM model, whereas RCME AP-HD showed significant effects of locomotor context and conversation.

Finally, regarding the smoothness of head movements, LDLJ V-HD showed only a significant Cohort $\times$ Context $\times$ Conversation interaction. LDLJ ML-HD did not show significant effects in the GLMM model. LDLJ AP-HD instead showed

significant Cohort $\times$ Context, Cohort $\times$ Conversation, and Cohort $\times$ Context $\times$ Conversation interactions, in the absence of significant main effects of cohort, locomotor context, or conversation.

4.4 Post-Hoc Comparisons

For the parameters which showed significant main effects or interactions in the GLMM models, post-hoc comparisons corrected using the False Discovery Rate (FDR) procedure were performed. All statistically significant comparisons are reported in Table 4.23.

Table 4.23: Significant model-based post-hoc comparisons after FDR correction.

Parameter	Comparison type	Comparison	Difference	q-value
Froude number	Context within cohort	Straight vs Unsignalised in OA, No Conversation	0.020	0.012
Froude number	Context within cohort	Signalised vs Unsignalised in OA, No Conversation	0.043	<0.001
Froude number	Context within cohort	Straight vs Unsignalised in OA, Conversation	0.013	0.044
Froude number	Context within cohort	Straight vs Signalised in YA, No Conversation	-0.041	<0.001
Froude number	Context within cohort	Signalised vs Unsignalised in YA, No Conversation	0.041	<0.001
Froude number	Conversation within cohort/context	No Conversation vs Conversation in OA, Signalised	0.044	0.002
Froude number	Conversation within cohort/context	No Conversation vs Conversation in YA, Signalised	0.035	0.002
stride speed	Context within cohort	Straight vs Unsignalised in OA, No Conversation	0.075	0.004
stride speed	Context within cohort	Signalised vs Unsignalised in OA, No Conversation	0.141	<0.001
stride speed	Context within cohort	Straight vs Signalised in OA, Conversation	0.077	0.017
stride speed	Context within cohort	Straight vs Unsignalised in OA, Conversation	0.050	0.017
stride speed	Context within cohort	Straight vs Signalised in YA, No Conversation	-0.120	<0.001
stride speed	Context within cohort	Signalised vs Unsignalised in YA, No Conversation	0.130	<0.001
stride speed	Conversation within cohort/context	No Conversation vs Conversation in OA, Signalised	0.150	<0.001
stride speed	Conversation within cohort/context	No Conversation vs Conversation in YA, Signalised	0.115	0.003
stride length	Context within cohort	Straight vs Unsignalised in OA, No Conversation	0.048	0.006
stride length	Context within cohort	Signalised vs Unsignalised in OA, No Conversation	0.080	0.003
stride length	Context within cohort	Straight vs Signalised in OA, Conversation	0.067	0.003
stride length	Context within cohort	Straight vs Unsignalised in OA, Conversation	0.041	0.004
stride length	Context within cohort	Straight vs Signalised in YA, No Conversation	-0.082	0.003
stride length	Context within cohort	Signalised vs Unsignalised in YA, No Conversation	0.071	0.003
stride length	Conversation within cohort/context	No Conversation vs Conversation in OA, Signalised	0.105	<0.001
stride length	Conversation within cohort/context	No Conversation vs Conversation in YA, Signalised	0.072	0.009
stride duration	Context within cohort	Straight vs Unsignalised in OA, No Conversation	-0.032	0.004

Continued on next page

Results

Parameter	Comparison type	Comparison	Difference	q-value
stride duration	Context within cohort	Signalised vs Unsignalised in OA, No Conversation	-0.055	<0.001
stride duration	Context within cohort	Straight vs Signalised in OA, Conversation	-0.030	0.042
stride duration	Context within cohort	Signalised vs Unsignalised in YA, No Conversation	-0.047	0.001
stride duration	Conversation within cohort/context	No Conversation vs Conversation in OA, Signalised	-0.054	0.003
stride duration	Conversation within cohort/context	No Conversation vs Conversation in YA, Signalised	-0.048	0.003
cadence	Context within cohort	Straight vs Unsignalised in OA, No Conversation	2.649	0.013
cadence	Context within cohort	Signalised vs Unsignalised in OA, No Conversation	5.336	<0.001
cadence	Context within cohort	Straight vs Signalised in YA, No Conversation	-3.077	0.031
cadence	Context within cohort	Signalised vs Unsignalised in YA, No Conversation	5.083	<0.001
cadence	Conversation within cohort/context	No Conversation vs Conversation in OA, Signalised	4.594	0.007
cadence	Conversation within cohort/context	No Conversation vs Conversation in YA, Signalised	4.306	0.006
aRMS/speed V-HD	Context within cohort	Straight vs Signalised in YA, Conversation	-0.393	0.002
aRMS/speed V-HD	Context within cohort	Straight vs Unsignalised in YA, Conversation	-0.239	0.001
aRMS/speed V-HD	Conversation within cohort/context	No Conversation vs Conversation in YA, Unsignalised	-0.222	0.037
aRMS/speed ML-HD	Context within cohort	Straight vs Unsignalised in OA, No Conversation	-0.189	0.004
aRMS/speed ML-HD	Context within cohort	Straight vs Unsignalised in YA, No Conversation	-0.235	0.002
aRMS/speed ML-HD	Context within cohort	Straight vs Signalised in YA, Conversation	-0.198	0.008
aRMS/speed ML-HD	Context within cohort	Straight vs Unsignalised in YA, Conversation	-0.117	0.008
aRMS/speed ML-HD	Conversation within cohort/context	No Conversation vs Conversation in OA, Straight	-0.148	0.026
aRMS/speed AP-HD	Context within cohort	Straight vs Unsignalised in OA, No Conversation	-0.158	0.040
RCME V-HD	Cohort within context	OA vs YA in Straight, No Conversation	0.225	<0.001
RCME V-HD	Cohort within context	OA vs YA in Straight, Conversation	0.170	0.004
RCME V-HD	Cohort within context	OA vs YA in Signalised, No Conversation	0.200	0.003
RCME V-HD	Cohort within context	OA vs YA in Signalised, Conversation	0.158	0.010
RCME V-HD	Cohort within context	OA vs YA in Unsignalised, No Conversation	0.177	0.004
RCME V-HD	Cohort within context	OA vs YA in Unsignalised, Conversation	0.183	0.003
RCME V-HD	Context within cohort	Straight vs Unsignalised in YA, No Conversation	-0.053	0.047
RCME V-HD	Conversation within cohort/context	No Conversation vs Conversation in YA, Straight	-0.052	0.008
RCME ML-HD	Context within cohort	Straight vs Signalised in OA, Conversation	0.087	0.004
RCME ML-HD	Context within cohort	Signalised vs Unsignalised in OA, Conversation	-0.098	0.002
RCME AP-HD	Context within cohort	Straight vs Unsignalised in OA, No Conversation	0.112	<0.001
RCME AP-HD	Context within cohort	Straight vs Signalised in OA, Conversation	0.085	0.037
RCME AP-HD	Context within cohort	Straight vs Unsignalised in OA, Conversation	0.057	0.034
RCME AP-HD	Context within cohort	Straight vs Signalised in YA, No Conversation	0.073	0.045
RCME AP-HD	Context within cohort	Straight vs Unsignalised in YA, No Conversation	0.081	0.013
RCME AP-HD	Context within cohort	Straight vs Signalised in YA, Conversation	0.105	0.001
RCME AP-HD	Context within cohort	Straight vs Unsignalised in YA, Conversation	0.051	0.012
RCME AP-HD	Conversation within cohort/context	No Conversation vs Conversation in OA, Unsignalised	-0.114	<0.001

Continued on next page

Results

Parameter	Comparison type	Comparison	Difference	q-value
LDLJ V-HD	Context within cohort	Straight vs Unsignalised in YA, No Conversation	0.170	0.016
LDLJ V-HD	Conversation within cohort/context	No Conversation vs Conversation in YA, Unsignalised	-0.120	0.023
LDLJ AP-HD	Context within cohort	Straight vs Signalised in YA, No Conversation	0.316	0.008
LDLJ AP-HD	Context within cohort	Straight vs Unsignalised in YA, No Conversation	0.353	<0.001
LDLJ AP-HD	Conversation within cohort/context	No Conversation vs Conversation in YA, Straight	0.280	0.001

Overall, the post-hoc comparisons showed that the significant effects observed in the GLMM models were mainly attributable to differences between the different locomotor contexts and, to a lesser extent to the differences between conditions with and without conversation. In addition, for some head-movement parameters, significant differences between the two cohorts emerged, particularly for RCME V-HD and LDLJ AP-HD.

4.5 Comparison between Aggregated and Context-Aware Analysis

In order to compare the results obtained with and without the inclusion of contextual information, the results of the aggregated analysis were compared with those derived from the context-aware analysis. Table 4.24 reports, for each parameter, the statistically significant effects observed in the two different approaches.

Table 4.24: Comparison between aggregated and context-aware analyses. Only statistically significant effects are reported.

Parameter	Aggregated analysis	Context-aware analysis
Froude number	No significant effect	Cohort effect; Context effect; Context $\times$ Conversation interaction
stride speed	No significant effect	Cohort effect; Context effect; Context $\times$ Conversation interaction
stride length	No significant effect	Context effect; Cohort $\times$ Context interaction; Context $\times$ Conversation interaction
stride duration	No significant effect	Context effect; Context $\times$ Conversation interaction
cadence	No significant effect	Context effect; Context $\times$ Conversation interaction
aRMS/speed V-HD	No significant effect	Cohort $\times$ Context $\times$ Conversation interaction
aRMS/speed ML-HD	No significant effect	Context effect; Conversation effect
aRMS/speed AP-HD	No significant effect	Context effect; Conversation effect
RCME V-HD	Cohort effect	Cohort effect; Cohort $\times$ Conversation interaction
RCME AP-HD	No significant effect	Context effect; Conversation effect
LDLJ V-HD	No significant effect	Cohort $\times$ Context $\times$ Conversation interaction
LDLJ AP-HD	No significant effect	Cohort $\times$ Context interaction; Cohort $\times$ Conversation interaction; Cohort $\times$ Context $\times$ Conversation interaction

For all spatiotemporal gait parameters, the aggregated analysis did not show significant cohort effects. In contrast, the context-aware analysis showed significant effects of locomotor context and of the Context $\times$ Conversation interaction for all the parameters considered. In addition, the Froude number and stride speed

also showed a significant cohort effect, whereas stride length showed a significant Cohort $\times$ Context interaction.

Regarding the parameters related to head movements, the aggregated analysis showed a significant cohort effect only for RCME V-HD. In the context-aware analysis, additional significant effects emerged. In particular, for the magnitude of head movements, aRMS/speed V-HD showed a significant Cohort $\times$ Context $\times$ Conversation interaction, whereas aRMS/speed ML-HD and aRMS/speed AP-HD showed significant effects of locomotor context and conversation.

For the regularity of head movements, RCME V-HD showed a significant cohort effect in both the aggregated and context-aware analyses. In the context-aware analysis, RCME V-HD also showed a significant Cohort $\times$ Conversation interaction. RCME AP-HD did not show significant effects in the aggregated analysis, whereas significant effects of locomotor context and conversation emerged in the context-aware analysis.

Finally, for the smoothness of head movements, the aggregated analysis did not show significant effects. In the context-aware analysis, LDLJ V-HD showed a significant Cohort $\times$ Context $\times$ Conversation interaction, whereas LDLJ AP-HD showed significant Cohort $\times$ Context, Cohort $\times$ Conversation, and Cohort $\times$ Context $\times$ Conversation interactions.

Overall, the comparison between the two approaches shows that the aggregated analysis highlighted a significant cohort effect only for RCME V-HD, whereas the context-aware analysis made it possible to identify additional significant effects related to locomotor context, conversation, and to the interactions among the factors considered.

Chapter 5

Discussion

5.1 Main findings

The present study investigated differences associated with locomotor context, the presence of conversation, and age in gait characteristics and head movements during real-world urban locomotion. The most innovative aspect of this work lies in the integration of biomechanical parameters, wearable sensors, and contextual information derived from egocentric videos, which enabled locomotor behaviour to be interpreted in relation to the situation in which it occurred.

The aggregated analysis, conducted without distinguishing between locomotor context and the presence of conversation, identified a significant difference between young and older adults exclusively for RCME V-HD. In particular, older adults showed higher values than young adults. Since higher RCME values indicate greater complexity and lower signal predictability, this finding suggests a different temporal organisation of vertical head movements in the two cohorts. However, this difference should not automatically be interpreted as impaired motor control, as greater entropy may reflect both lower regularity and a more variable and adaptive strategy.

By contrast, the aggregated analysis provided no statistically significant evidence of differences between cohorts in spatiotemporal gait parameters, head-movement amplitude, the other RCME components, or smoothness. This finding suggests that, in the sample of healthy older adults considered, overall locomotor performance was largely preserved and that age-related differences did not manifest as generalised alterations across all the parameters analysed.

The context-aware analysis revealed a more complex picture. All spatiotemporal gait parameters were influenced by locomotor context and by the Context $\times$ Conversation interaction. FDR-corrected post-hoc comparisons showed that, particularly in the absence of conversation, signalised crossings were characterised by higher

stride speed, stride length, cadence, and Froude number, together with shorter stride duration, than unsignalised crossings. During signalised crossings, however, the presence of conversation was associated in both cohorts with reductions in stride speed, stride length, cadence, and Froude number, as well as an increase in stride duration. Conversation therefore did not exert a uniform effect on gait, but primarily modified the strategy adopted in the signalised-crossing context.

Froude number and stride speed also showed a significant main effect of cohort. The descriptive values and model estimates indicated a tendency for older adults to show higher values than young adults. However, individual between-cohort comparisons within specific combinations of context and conversation were not significant after FDR correction. The cohort effect should therefore be interpreted as an overall difference within the model, rather than as evidence of a robust difference in every individual condition. Stride length, by contrast, showed a Cohort $\times$ Context interaction, indicating that the relationship between the two cohorts varied according to locomotor context, without identifying a consistent superiority of either cohort across all conditions.

Head movements also showed specific adaptations. Normalised head-acceleration amplitude was generally greater during crossings than during straight walking in certain directions and conditions, suggesting increased head movement relative to progression speed. This increase may reflect both greater demands for environmental exploration and more pronounced postural adaptations and cannot, in isolation, be regarded as an indicator of instability.

For RCME, the vertical component represented the most robust age-related finding, with higher values in older adults across all combinations of context and conversation. The anteroposterior component, by contrast, was mainly modulated by context and conversation. For smoothness, measured using LDLJ, effects emerged primarily as interactions. In young adults, straight walking without conversation was generally characterised by smoother head movements than crossings, whereas conversation produced different effects depending on the direction and context considered.

Overall, the findings indicate that urban locomotion cannot be adequately described using a single aggregated value. Environmental context and the presence of conversation systematically modify locomotor behaviour and head control. The context-aware approach therefore made it possible to distinguish specific adaptations that may be attenuated or absorbed into residual variability when heterogeneous conditions are analysed together.

5.2 Effects of locomotor context on gait parameters

The context-aware analysis showed that all spatiotemporal gait parameters were modulated by locomotor context. This finding confirms that, in real-world environments, gait is not an invariant behaviour, but is continuously adapted to the perceptual, decisional, and motor demands of the surrounding situation.

The clearest pattern emerged during signalised crossings in the absence of conversation. Under this condition, both groups generally adopted a faster locomotor strategy than during unsignalised crossings, characterised by higher stride speed, a higher Froude number, greater cadence, and shorter stride duration. Stride length was also greater in several comparisons. This behaviour may be interpreted as a response to the presence of an external temporal constraint. The traffic signal defines a time window within which the crossing must be completed and may therefore encourage an acceleration of progression once the crossing has begun.

The Froude number provides further insight into this interpretation, as it expresses speed normalised to lower-limb length. The differences between contexts observed for this parameter therefore indicate that the increase in speed during signalised crossings was not solely attributable to participants' anthropometric characteristics, but reflected a genuine modification of locomotor strategy.

Unsignalised crossings impose different demands. In the absence of an external signal indicating when to cross, pedestrians must independently assess traffic, vehicle distance and speed, and the availability of a safe gap. Once the crossing has begun, behaviour may also depend on the prevailing traffic conditions. The greater variability observed in this condition is therefore compatible with a less standardised strategy that is more strongly dependent on the specific characteristics of each event.

A particularly relevant finding concerns the Context $\times$ Conversation interaction, which was significant for all spatiotemporal parameters. During signalised crossings, conversation was associated in both cohorts with reductions in stride speed, stride length, cadence, and Froude number, together with an increase in stride duration. Conversation therefore appears to attenuate the acceleration strategy observed in the same condition when participants were not engaged in social interaction.

This pattern is consistent with the sharing of attentional resources between locomotion, environmental monitoring, and conversation. Signalised crossings require the integration of the traffic signal, surrounding traffic, and the time available, while conversation simultaneously requires language processing and interaction with the conversational partner. The reduction in progression rate may therefore represent a strategy adopted to maintain control of the task in the presence of concurrent demands. However, because the conversation was face-to-face rather

than a standardised cognitive task, the finding cannot be attributed exclusively to cognitive load. Social regulation of walking speed and orientation towards the conversational partner may also have contributed to the observed changes.

The absence of a uniform main effect of conversation, together with the presence of the Context $\times$ Conversation interaction, is a substantive finding. It indicates that conversation does not necessarily cause slowing in every situation; rather, its effect depends on the context in which it occurs. This feature would have been difficult to detect through a global comparison between walking with and without conversation.

With regard to age-related differences, Froude number and stride speed showed a main effect of cohort, with values tending overall to be higher in older adults. This finding does not correspond to the typical pattern of reduced walking speed often associated with ageing. It may be explained, at least in part, by the characteristics of the sample, which consisted of healthy and independent older adults, and by the ecological nature of the protocol, in which behaviour was influenced by real-world conditions and interaction with the environment. Moreover, no OA–YA comparison within an individual combination of context and conversation remained significant after FDR correction. It is therefore not appropriate to conclude that older adults walked systematically faster than young adults across all conditions.

The significant Cohort $\times$ Context interaction observed for stride length instead indicates that the difference between cohorts was not constant across locomotor contexts. This finding suggests the presence of partly different adaptation strategies in young and older adults, but the absence of significant simple comparisons after correction for multiple testing requires a cautious interpretation. Overall, locomotor context and its interaction with conversation were more evident determinants of gait parameters than chronological age considered in isolation.

5.3 Effects of locomotor context and conversation on head movements

Head movements showed different patterns according to the metric and anatomical direction considered. This finding confirms that amplitude, temporal complexity, and smoothness describe distinct aspects of head control and cannot be interpreted as interchangeable measures.

5.3.1 Head-movement amplitude

The aRMS normalised to speed quantifies the intensity of head accelerations while accounting for progression speed. A higher value therefore indicates that, relative

to walking speed, the head undergoes oscillations or changes in acceleration of greater amplitude.

The vertical component showed a significant Cohort $\times$ Context $\times$ Conversation interaction. Post-hoc comparisons indicated that, in young adults in the presence of conversation, signalised and unsignalised crossings were associated with higher values than straight walking. Furthermore, during unsignalised crossings in young adults, conversation was associated with an increase in vertical aRMS/speed. This pattern suggests that the combination of crossing and social interaction modified the intensity of vertical head movements particularly in young adults.

For the mediolateral component, unsignalised crossings showed higher values than straight walking in both cohorts in the absence of conversation. In young adults during conversation, both signalised and unsignalised crossings also showed higher values than straight walking. In older adults, during straight walking, conversation was associated with an increase in mediolateral amplitude.

These findings are compatible with greater lateral exploration of the environment during crossings, which is required to monitor vehicles and other road users. However, the inertial sensor measures head movement rather than gaze direction directly. It is therefore not possible to distinguish with certainty between movements aimed at visual exploration and oscillations related to postural control or gait dynamics.

For the anteroposterior component, older adults showed a higher value during unsignalised crossings than during straight walking in the absence of conversation. This difference may reflect greater sagittal-plane adjustments of the head during a condition in which the independent assessment of the environment plays a central role.

Overall, the increase in aRMS/speed during crossings should not automatically be interpreted as reduced stability. In an urban environment, greater head-movement amplitude may also represent a functional response to the demands of exploration and monitoring. Integration with eye-tracking measures would be required to distinguish more clearly between intentional visual exploration and reduced head stabilisation.

5.3.2 Complexity and regularity of head movements

RCME quantifies the temporal complexity and predictability of a signal. Higher values indicate less predictable and more complex dynamics, whereas lower values describe a more repetitive and regular pattern. RCME should therefore not be interpreted according to a simple relationship in which a higher value is necessarily better or worse.

RCME V-HD represented the most robust cohort-related finding. Older adults showed higher values than young adults across all the combinations of context and

conversation considered. The persistence of this difference in both the aggregated analysis and the context-aware comparisons suggests that greater complexity of vertical head movements was a relatively stable characteristic of the older cohort in the sample examined.

One possible interpretation is that older adults employed a more variable and less predictable vertical head-control strategy. This behaviour may reflect a reduced ability to maintain a repetitive pattern, a greater need to make postural corrections, or an adaptive strategy aimed at responding to environmental variation. The available data do not establish which of these explanations predominates; the finding should therefore be described as a different organisation of movement rather than directly as a deficit.

In young adults, RCME V-HD was also higher during unsignalised crossings than during straight walking in the absence of conversation. During straight walking, conversation was associated with an increase in vertical RCME. This suggests that, in young adults, both environmental complexity and social interaction may make vertical head movement less predictable.

RCME AP-HD showed effects of context and conversation. In both cohorts, straight walking generally showed higher values than crossings. This finding suggests that, during crossings, anteroposterior head movement may become more constrained and repetitive, possibly in order to maintain a more controlled orientation while monitoring the road and carrying out the crossing. In older adults, during unsignalised crossings, conversation was instead associated with an increase in anteroposterior complexity, indicating that social interaction modified the temporal organisation of movement in this specific condition.

No overall significant effects emerged for RCME ML-HD in the GLMM. Although some isolated post-hoc comparisons reached significance, the absence of a global effect requires caution and does not support the existence of a generalised mediolateral modulation.

5.3.3 Smoothness of head movements

Smoothness was assessed using LDLJ. Since more negative values indicate a greater presence of abrupt changes in acceleration, less negative values correspond to smoother and more continuous movements.

A Cohort $\times$ Context $\times$ Conversation interaction emerged for LDLJ V-HD. In young adults in the absence of conversation, straight walking was characterised by greater smoothness than unsignalised crossings. In the same group, during unsignalised crossings, conversation was associated with less negative values and therefore with smoother vertical movements. This finding shows that conversation does not necessarily impair smoothness, but may modify movement organisation differently depending on the situation.

For LDLJ AP-HD, Cohort $\times$ Context, Cohort $\times$ Conversation, and Cohort $\times$ Context $\times$ Conversation interactions emerged. However, the significant post-hoc comparisons were concentrated in young adults. In the absence of conversation, straight walking was smoother than both signalised and unsignalised crossings. During straight walking, conversation was instead associated with reduced antero-posterior smoothness. The concurrent task therefore appears to have introduced more abrupt movements into a condition that was normally regular, whereas its effect could differ in other directions and contexts.

No significant effects emerged for LDLJ ML-HD. Overall, the smoothness findings do not support a simple global distinction between young and older adults. Rather, they show that the fluidity of head movements varies according to the specific combination of anatomical direction, locomotor context, conversation, and cohort.

5.4 Age-related differences in real-world locomotion

In the present study, age-related differences did not manifest as a generalised impairment of gait in older adults. In the aggregated analysis, no spatiotemporal parameter differed significantly between cohorts. This finding is plausibly related to the characteristics of the older participants, who were healthy, independent, and free from relevant neurological or musculoskeletal disorders. In a sample of this kind, conventional gait parameters may remain within functional ranges similar to those observed in young adults.

However, the absence of statistical significance does not demonstrate equivalence between cohorts. The number of participants was small, and inter-individual variability may have limited the ability to detect effects of modest magnitude. The large number of events increases the number of available observations, but it does not replace the number of independent participants when generalising age-related differences.

In the context-aware analysis, Froude number and stride speed showed a main effect of cohort. The tendency towards higher values in older adults should be interpreted cautiously, as no significant OA–YA comparisons emerged within individual conditions after FDR correction. The finding may reflect an overall difference between cohorts in the model, but does not demonstrate that older adults walked faster in every context. Furthermore, the selection of healthy and active older adults and the specific characteristics of the route may have produced a sample that was not representative of the wider older population.

Stride length showed a Cohort $\times$ Context interaction, suggesting that the two cohorts adapted this parameter in ways that were not entirely overlapping across contexts. Again, the absence of specific between-cohort comparisons that remained

significant after FDR correction does not allow the difference to be localised confidently to a single condition.

The most consistent age-related finding was observed for RCME V-HD. Older adults showed greater complexity and lower predictability of vertical head movements in all the conditions considered. The fact that this difference emerged even when spatiotemporal parameters were largely preserved suggests that head control may be more sensitive than conventional gait descriptors in identifying age-related adaptations.

One possible interpretation is that older adults were able to maintain overall locomotor performance comparable to that of young adults by using different head-control strategies. Greater vertical complexity may reflect an increase in postural corrections, different transmission of body oscillations to the head, or more variable exploratory behaviour. However, without direct measures of stability, balance, and visual behaviour, it is not possible to determine whether this pattern represents effective compensation or an indication of reduced control.

The interactions observed for LDLJ V-HD and LDLJ AP-HD further indicate that any age-related differences in smoothness were not constant, but depended on context and conversation. Since the significant post-hoc comparisons occurred predominantly within the young-adult cohort, it would be inappropriate to conclude that older adults had globally less smooth movements. The main finding is instead that the way in which smoothness is modulated by the experimental conditions may differ between cohorts.

Overall, the data suggest a selective rather than generalised pattern of locomotor ageing. Healthy older adults maintained relatively preserved spatiotemporal parameters, whereas more consistent differences emerged in the temporal organisation of head movements. These findings support the inclusion of head-control measures in the assessment of real-world mobility, while avoiding the automatic attribution of pathological significance to such differences.

5.5 Importance of the context-aware approach

The comparison between the aggregated and context-aware analyses represents one of the principal contributions of the present study. The aggregated analysis identified only a cohort difference in RCME V-HD, whereas the inclusion of locomotor context and conversation revealed numerous adaptations in gait and head movements.

This finding does not imply that the context-aware analysis artificially generated new effects. On the contrary, aggregation combines observations from conditions characterised by different demands. If a parameter increases in one context and decreases in another, or if conversation produces an effect only during signalised

crossings, the aggregated value may attenuate or cancel out these differences. In this way, context-related variability is incorporated into residual variability rather than being used as explanatory information.

The Context $\times$ Conversation interaction observed for all spatiotemporal parameters provides a clear example of this phenomenon. Conversation reduced the rate of progression particularly during signalised crossings, whereas the same pattern was not observed in the other contexts. An aggregated comparison between conditions with and without conversation would have combined effects of different magnitudes and, in some cases, different directions.

Head-movement parameters also showed marked directional and contextual specificity. Amplitude, complexity, and smoothness were not modified in the same way along the vertical, mediolateral, and anteroposterior axes. The context-aware analysis therefore made it possible to recognise that different components of head control respond to different environmental and behavioural demands.

It is nevertheless important to distinguish contextual effects from age-related differences. Most of the effects emerging from the context-aware analysis concerned locomotor context, conversation, or their interactions, rather than a direct difference between young and older adults. The context-aware approach therefore revealed primarily situational adaptations in locomotion, while cohort differences were more selective and particularly robust for RCME V-HD.

The use of GLMMs made it possible to account for the dependence between observations obtained from the same participant and for the imbalance in the number of events across conditions. Post-hoc comparisons corrected using FDR also limited the risk of interpreting results as significant merely because of the large number of comparisons. Nevertheless, complex effects must still be interpreted cautiously, particularly when a significant global interaction is not accompanied by significant simple comparisons after correction for multiple testing.

Overall, contextual information is not an ancillary variable, but a necessary component for understanding the meaning of parameters acquired in real-world settings. The same change in speed or head movement may reflect a functional adaptation to environmental demands rather than a reduction in locomotor capacity. The context-aware approach therefore enables a more ecological and specific characterisation of everyday mobility, distinguishing relatively stable individual characteristics from adaptations produced by the situations encountered.

5.6 Performance of the Automatic Context-Recognition Pipeline

A central element of the present study is represented by the pipeline developed to associate biomechanical parameters with the locomotor context observed in

egocentric videos. The use of pre-trained networks made it possible to semi-automatically identify different locomotor situations, such as pedestrian crossings and the presence of traffic lights, enabling the subsequent context-aware analysis.

For pedestrian crossing recognition using the Places365-based network, the application of temporal post-processing improved performance compared with frame-by-frame classification. At frame level, accuracy increased from 0.871 to 0.900, recall from 0.593 to 0.716, and the F1-score from 0.705 to 0.788, while precision remained substantially stable. This indicates that temporal information is useful to stabilise the network output and to reduce fluctuations due to the instantaneous variability of the egocentric scene.

An improvement also emerged at event level. Using a bidirectional overlap criterion of 20%, the post-processed output correctly identified 265 of the 354 manually annotated crossings, with a recall of 0.749. Compared with the RAW output, precision increased from 0.029 to 0.365 and the F1-score from 0.054 to 0.490. In addition, the number of missed events decreased from 176 to 89, whereas the number of extra events decreased from 6022 to 462. The reduction in the fragmentation index from 41 to 32 further suggests that the temporal stabilization procedures, based on EMA and hysteresis, improved the temporal coherence of the automatic detections. However, the persistence of residual fragmentation indicates that maintaining a stable classification over time remains a critical issue for pre-trained models applied to egocentric videos acquired in real-world environments.

Despite these improvements, the event-level performance of Places365 still highlights some limitations. The precision value of 0.365 indicates the residual presence of false positives, whereas the reduction in recall and F1-score as the increased overlap threshold suggests that the temporal localisation of events was not always precise. Therefore, the Places365-based module proved useful to support the identification of crossings, but it was not sufficient to completely replace manual review.

Regarding the recognition of signalised crossings, the YOLOv11-based module showed more robust performance at event level than at frame-by-frame level. At frame level, the classification achieved an accuracy of 0.931, a precision of 0.943, a recall of 0.630, and an F1-score of 0.755. The high precision indicates that, when the traffic lights were detected, the prediction was generally reliable; the lower recall instead suggests that the traffic lights were not detected in all frames in which it was present.

At event level, considering a crossing as signalised when at least 20% of its frames contained a traffic-light detection, the module correctly identified 81 of the 91 signalised crossings, with a recall of 0.890, a precision of 0.771, and an F1-score of 0.827. The aggregation of predictions at event level therefore partially compensates for missed detections at frame level and provides a classification useful to distinguish between signalised and unsignalised crossings.

Overall, the pre-trained networks, combined with post-processing strategies and temporal aggregation, represented a useful tool to support the annotation of the locomotor context. However, the observed performance does not allow the process to be considered fully automatic. For this reason, the automatically generated labels were manually reviewed before the statistical analysis, reducing the risk that context-classification errors could propagate to the biomechanical results.

A relevant limitation of the pipeline is represented by the use of pre-trained models which were not specifically optimised on egocentric videos acquired during urban locomotion protocols. Images acquired through smart glasses are characterised by continuous camera motion, rapid changes in viewpoint, partial occlusions, changes in lighting conditions, and differences in the distance of relevant objects. These aspects may explain part of the observed errors and highlight the need to develop specifically annotated datasets for fine-tuning or retrain the networks in real-world first-person locomotion scenarios.

5.7 Methodological Considerations

A relevant methodological aspect of the present study concerns the choice to analyse the data at the event level. The aggregation of biomechanical parameters within each event made it possible to associate each observation with a specific combination of cohort, locomotor context, and conversation condition. This approach is particularly useful in real-world protocols, in which event duration and the number of available strides may vary in different subjects and conditions.

The use of mixed-effect models also made it possible to account for the hierarchical structure of the data by including the subject as a random effect. This aspect is important because multiple events derive from the same participant and therefore could not be considered completely independent observations.

The combination of contextual annotation and event-level analysis made it possible to interpret gait and head-movement parameters in relation to the specific locomotor situation in which they were acquired. In this way, variability between events was not treated only as a residual variability, but as information related to the different environmental and behavioural demands.

5.8 Study Limitations

The results of the present study should be interpreted in light of some limitations.

Firstly, the number of participants included in the analysis was relatively small. Although the experimental protocol generated a large number of locomotor events and strides, the sample consisted of a limited number of younger adults and older adults. This aspect may limit the generalisability of the results and reduce the

ability to identify more subtle cohort-related effects, especially for parameters characterised by high inter-subject variability.

Secondly, the present study involved only healthy older adults who were autonomous and free from neurological or musculoskeletal disorders capable of significantly impairing locomotion. Consequently, the age-related differences observed in this sample may be less marked than those expected in clinical populations, frail subjects, or older adults with impairments in balance and motor control.

A further limitation concerns the experimental environment, which consisted of a single urban route. Although the protocol was designed to include different locomotor contexts, it does not necessarily represent the full variability of situations encountered in daily life. Other scenarios, such as crowded environments, uneven surfaces, stairs, obstacles, or more complex intersections, could induce different locomotor strategies.

From a methodological point of view, the automatic context recognition pipeline required a manual review phase of the labels. This aspect represents a limitation of the proposed approach, as it indicates that the performance of the pre-trained networks and post-processing procedures is not sufficient yet to obtain a fully automatic process.

In addition, the analysis was conducted at the event level, using mean values of the biomechanical parameters calculated within each locomotor event. This approach made it possible to associate each observation with a specific context and a specific conversation condition, but it does not allow a detailed description of intra-event variations or transitions between different locomotor phases.

Finally, the present study focused on spatiotemporal gait parameters and head movements, without directly including measures of visual behaviour. The integration of eye-tracking data could provide additional information on the perceptual and attentional mechanisms underlying the observed locomotor adaptations.

Despite these limitations, the present work demonstrates the feasibility of a context-aware approach for the characterisation of real-world locomotion and provides preliminary evidence on the importance of integrating contextual information with the analysis of gait parameters and head movements.

5.9 Future Perspectives

The results of the present study open several perspectives for future applications and methodological developments.

Firstly, it will be necessary to validate the proposed approach on larger samples characterised by greater heterogeneity in terms of age, functional level, and health status. The inclusion of a larger number of participants would increase the statistical power of the analysis and allow the effects of cohort, locomotor context, and

conversation observed in the present study to be verified more robustly.

Secondly, the application of context-aware analysis to clinical populations could represent a particularly relevant development. Subjects with neurological disorders, balance impairments, frailty, or increased risk of falling may show more marked locomotor alterations or different adaptation strategies compared with the healthy older adults included in the present work. In these populations, contextual information could allow the identification of alterations which would not emerge from the aggregated analyses of gait parameters.

Further studies should also expand the number and type of locomotor contexts considered. In addition to straight walking and pedestrian crossings, more complex scenarios could be included, such as crowded environments, stairs, obstacles, uneven surfaces, or more articulated urban intersections.

From a methodological point of view, a future objective concerns the improvement of the automatic context recognition pipeline. The construction of a specific annotated dataset, composed of egocentric videos acquired during urban locomotion protocols, could allow fine-tuning or retraining of the networks used, adapting them more effectively to the characteristics of recordings acquired through smart glasses.

A further methodological development concerns the automatic recognition of a traffic light state. The annotation of a specific dataset containing traffic lights observed from an egocentric perspective and classified according to their light state could allow the training of models so they would be able to directly recognise not only the presence of traffic lights, but also the colour or functional state. This would make it possible to distinguish more robustly between conditions such as green, amber, and red traffic lights, reducing the dependence on colour-based rules and manual review.

In addition, future developments could integrate temporal models or more robust tracking strategies, which would be able to exploit the continuity between consecutive frames and reduce the fragmentation of automatically recognised events. These approaches could be applied both to crossing recognition and to traffic light state classification, improving the temporal coherence of the automatically generated labels.

The integration of eye-tracking measures could provide additional information on the visual behaviour of participants. The joint analysis of eye movements, head movements, and gait parameters would allow a more complete description of the perceptual-motor mechanisms involved in urban locomotion, especially on pedestrian crossings and in conditions of greater environmental complexity.

Finally, the integration of wearable sensors, smart glasses, and automatic context recognition algorithms could contribute to the development of context-aware digital biomarkers, capable of describing not only the global motor performance of individuals, but also their ability to adapt locomotor behaviour to the different environmental and cognitive conditions in their daily life.

Chapter 6

Conclusion

The present study proposed a context-aware approach for the characterisation of real-world locomotion, based on the integration of wearable sensors and egocentric videos acquired using smart glasses. This approach made it possible to associate gait and head-movement parameters with the locomotor contexts encountered during urban locomotion, distinguishing between straight walking, signalised crossings, and unsignalised crossings, in the presence or absence of conversation.

The results showed that locomotor context and the presence of conversation significantly influence locomotor behaviour in real-world conditions. In particular, all spatiotemporal gait parameters showed significant effects of context and of the Context $\times$ Conversation interaction in the context-aware analysis, whereas no significant differences had emerged for these parameters in the aggregated analysis.

Head movements also showed a selective dependence on the factors considered. Among the analyzed parameters, RCME V-HD was the indicator most clearly associated with differences between younger adults and older adults, showing a significant cohort effect in both the aggregated and context-aware analyses. Other parameters, such as aRMS/speed V-HD, LDLJ V-HD, and LDLJ AP-HD, instead showed significant effects mainly in the form of interactions, indicating that differences in head movements may depend on the specific combination of cohort, locomotor context, and presence of conversation.

Overall, the results of the present work show that the explicit consideration of locomotor context represents a fundamental element to understand locomotion in real-world environments. Aggregating data from heterogeneous locomotor situations may in fact mask specific adaptations, reducing the ability to identify effects related to context, conversation, and the interactions among the factors considered.

Finally, the present study highlights the potential of integrating wearable sensors, smart glasses, and automatic context recognition techniques for the development of innovative approaches to the assessment of daily-life mobility. Although the proposed pipeline still requires a manual review phase, it represents a step toward

locomotion analyses which are more ecologically valid and sensitive to the real conditions in which movement occurs. Further refinement of the automatic recognition procedures, also through fine-tuning of the networks on specific egocentric datasets, could reduce the need for manual intervention and make the approach more applicable to large-scale studies or prolonged real-world monitoring. In perspective, these methodologies could contribute to the development of context-aware digital biomarkers and support the functional assessment and monitoring of older and clinical populations.

Bibliography

- [1] Stacy Fritz and Michelle Lusardi. «White paper: “walking speed: the sixth vital sign”». In: *Journal of geriatric physical therapy* 32.2 (2009), pp. 2–5 (cit. on pp. 1, 8).
- [2] Yi Ma, Zhipan Niu, Meng Shi, Wei Xie, Zuoan Hu, Yidong Wei, Tian Zeng, and Eric Wai Ming Lee. «Unraveling human crowd dynamics through the foot tracking of pedestrians». In: *Science Advances* 11.26 (2025), eadw2688 (cit. on p. 2).
- [3] Jennifer A. Oxley, Elfriede Ihlen, Brian N. Fildes, Judith L. Charlton, and Ross H. Day. «Crossing roads safely: An experimental study of age differences in gap selection by pedestrians». In: *Accident Analysis & Prevention* 37.5 (2005), pp. 962–971. ISSN: 0001-4575. DOI: <https://doi.org/10.1016/j.aap.2005.04.017>. URL: <https://www.sciencedirect.com/science/article/pii/S0001457505000795> (cit. on p. 2).
- [4] Mark B. Neider, John G. Gaspar, Jason S. McCarley, James A. Crowell, Henry Kaczmarski, and Arthur F. Kramer. «Walking and talking: Dual-task effects on street crossing behavior in older adults.» In: *Psychology and Aging* 26.2 (2011), pp. 260–268. ISSN: 1939-1498(Electronic),0882-7974(Print). DOI: 10.1037/a0021566 (cit. on p. 2).
- [5] Charlotte Hennah and Michail Doulas. «Dual-task walking on real-world surfaces: Adaptive changes in walking speed, step width and step height in young and older adults». In: *Experimental Gerontology* 177 (2023), p. 112200. ISSN: 0531-5565. DOI: 10.1016/j.exger.2023.112200. URL: <https://www.sciencedirect.com/science/article/pii/S0531556523001213> (cit. on p. 2).
- [6] Victoria I. Nicholls, Jan Wiener, Andrew Isaac Meso, and Sebastien Miellet. «The impact of perceptual complexity on road crossing decisions in younger and older adults». In: *Scientific Reports* 14.1 (Jan. 2024), p. 479. ISSN: 2045-2322. DOI: 10.1038/s41598-023-49456-9. URL: <https://doi.org/10.1038/s41598-023-49456-9> (cit. on p. 2).

- [7] G. A. Zito, D. Cazzoli, L. Scheffler, M. Jäger, R. M. Müri, U. P. Mosimann, T. Nyffeler, F. W. Mast, and T. Nef. «Street crossing behavior in younger and older pedestrians: an eye- and head-tracking study». In: *BMC Geriatrics* 15.1 (Dec. 2015), p. 176. ISSN: 1471-2318. DOI: 10.1186/s12877-015-0175-0. URL: <https://doi.org/10.1186/s12877-015-0175-0> (cit. on pp. 2, 11).
- [8] Aurélie Dommès and Viola Cavallo. «The role of perceptual, cognitive, and motor abilities in street-crossing decisions of young and older pedestrians». In: *Ophthalmic and Physiological Optics* 31.3 (2011), pp. 292–301. DOI: <https://doi.org/10.1111/j.1475-1313.2011.00835.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-1313.2011.00835.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-1313.2011.00835.x> (cit. on p. 2).
- [9] Donald A Neumann and Elisabeth Roen Kelly. *Kinesiology of the musculoskeletal system: foundations for rehabilitation*. Mosby/Elsevier, 2010 (cit. on pp. 4–7).
- [10] ResearchGate. *Phases in human gait*. Accessed: 2026-05-16. 2018. URL: https://www.researchgate.net/figure/Phases-in-human-gait-The-swing-phase-as-the-name-suggests-is-defined-from-the-moment_fig1_326901491 (cit. on p. 6).
- [11] Michael W Whittle. *Gait analysis: an introduction*. Butterworth-Heinemann, 2014 (cit. on pp. 6, 7).
- [12] Pamela K Levangie and Cynthia C Norkin. *Joint structure and function: a comprehensive analysis*. FA Davis, 2011 (cit. on pp. 6, 7).
- [13] Physiopedia. *The Gait Cycle — Physiopedia*. [Online; accessed 16-May-2026]. 2025. URL: https://www.physio-pedia.com/index.php?title=The_Gait_Cycle&oldid=365136 (cit. on p. 7).
- [14] M Burnfield. «Gait analysis: normal and pathological function». In: *Journal of Sports Science and Medicine* 9.2 (2010), p. 353 (cit. on pp. 8–10).
- [15] Richard G. Stefanacci and Jayne R. Wilkinson. *Gait Disorders in Older Adults*. Web Page. 2025/08 2025. URL: <https://www.msmanuals.com/professional/geriatrics/gait-disorders-in-older-adults/gait-disorders-in-older-adults> (cit. on p. 8).
- [16] Thibault Warlop, Christine Detrembleur, Benjamin Bollens, Gaetan Stoquart, Frederic Crevecoeur, Anne Jeanjean, and Thierry M Lejeune. «Temporal organization of stride duration variability as a marker of gait instability in Parkinson’s disease.» In: *Journal of rehabilitation medicine* 48.10 (2016), pp. 865–871 (cit. on p. 8).

- [17] Brian R Umberger. «Stance and swing phase costs in human walking». In: *Journal of the Royal Society Interface* 7.50 (2010), pp. 1329–1340 (cit. on p. 8).
- [18] Ronita L Cromwell, Tina K Aadland-Monahan, Amy T Nelson, Sara M Stern-Sylvestre, and Bryan Seder. «Sagittal plane analysis of head, neck, and trunk kinematics and electromyographic activity during locomotion». In: *Journal of Orthopaedic & Sports Physical Therapy* 31.5 (2001), pp. 255–262 (cit. on pp. 9, 10).
- [19] JJ Kavanagh, Steven Morrison, and RS Barrett. «Coordination of head and trunk accelerations during walking». In: *European journal of applied physiology* 94.4 (2005), pp. 468–475 (cit. on p. 9).
- [20] Peter Hausamann, Martin Daumer, Paul R MacNeilage, and Stefan Glasauer. «Ecological momentary assessment of head motion: toward normative data of head stabilization». In: *Frontiers in human neuroscience* 13 (2019), p. 179 (cit. on pp. 9, 10).
- [21] Claudia Mazzà, Marco Iosa, Fabrizio Pecoraro, and Aurelio Cappozzo. «Control of the upper body accelerations in young and elderly women during level walking». In: *Journal of neuroengineering and rehabilitation* 5.1 (2008), p. 30 (cit. on p. 10).
- [22] JL Bellanca, KA Lowry, JM Vanswearingen, JS Brach, and MS Redfern. «Harmonic ratios: a quantification of step to step symmetry». In: *Journal of biomechanics* 46.4 (2013), pp. 828–831 (cit. on p. 10).
- [23] Aftab E Patla and Sharon E Walt. «Biomechanical Walking Pattern Changes in the Fit and Healthy Elderly». In: *Phys Ther* 70 (1990), pp. 340–347 (cit. on pp. 10, 11).
- [24] Stephanie Studenski et al. «Gait speed and survival in older adults». In: *Jama* 305.1 (2011), pp. 50–58 (cit. on p. 11).
- [25] Hanatsu Nagano, Rezaul K Begg, William A Sparrow, and Simon Taylor. «A comparison of treadmill and overground walking effects on step cycle asymmetry in young and older individuals». In: *Journal of applied biomechanics* 29.2 (2013), pp. 188–193 (cit. on p. 11).
- [26] Tal Krasovsky, Patrice L Weiss, and Rachel Kizony. «Older adults pay an additional cost when texting and walking: effects of age, environment, and use of mixed reality on dual-task performance». In: *Physical therapy* 98.7 (2018), pp. 549–559 (cit. on pp. 11, 12).

- [27] Jona Mustafovska, Peter H Wilson, Michael H Cole, and Thomas B McGuckian. «Locomotor-cognitive dual-tasking is reduced in older adults relative to younger: A systematic review with meta-analysis». In: *Gait & Posture* 120 (2025), pp. 177–191 (cit. on pp. 11, 15).
- [28] Grace A Gabriel, Laurence R Harris, Joshua J Gnanasegaram, Sharon L Cushing, Karen A Gordon, Bruce C Haycock, and Jennifer L Campos. «Age-related changes to vestibular heave and pitch perception and associations with postural control». In: *Scientific Reports* 12.1 (2022), p. 6426 (cit. on p. 11).
- [29] Amy Morris, Corinne Masegian, Angela Zhang, Madelyn Carrillo, Brian Székely, Nicholas G Murray, and Fang Jiang. «The effect of age and proprioceptive illusion susceptibility on gait». In: *Physiology & behavior* 287 (2024), p. 114717 (cit. on p. 11).
- [30] Atalie C Thompson et al. «Association between contrast sensitivity and physical function in cognitively healthy older adults: The brain networks and mobility function study». In: *The Journals of Gerontology: Series A* 78.8 (2023), pp. 1513–1521 (cit. on p. 11).
- [31] Kate Wilmut and Catherine Purcell. «Why are older adults more at risk as pedestrians? A systematic review». In: *Human factors* 64.8 (2022), pp. 1269–1291 (cit. on pp. 11–13).
- [32] John M Franchak, Brianna McGee, and Gabrielle Blanch. «Adapting the coordination of eyes and head to differences in task and environment during fully-mobile visual exploration». In: *PLoS one* 16.8 (2021), e0256463 (cit. on pp. 12, 13).
- [33] Anup Doshi and Mohan M Trivedi. «Head and eye gaze dynamics during visual attention shifts in complex environments». In: *Journal of vision* 12.2 (2012), pp. 9–9 (cit. on p. 12).
- [34] Aurélie Dommes. «Street-crossing workload in young and older pedestrians». In: *Accident Analysis & Prevention* 128 (2019), pp. 175–184 (cit. on pp. 12, 13).
- [35] Giuseppe Angelo Zito, Dario Cazzoli, Loreen Scheffler, Michael Jäger, René Martin Müri, Urs Peter Mosimann, Thomas Nyffeler, Fred W Mast, and Tobias Nef. «Street crossing behavior in younger and older pedestrians: an eye-and head-tracking study». In: *BMC geriatrics* 15.1 (2015), p. 176 (cit. on pp. 12–14).
- [36] Hagai Tapiro, Tal Oron-Gilad, and Yisrael Parmet. «Pedestrian distraction: The effects of road environment complexity and age on pedestrian’s visual attention and crossing behavior». In: *Journal of safety research* 72 (2020), pp. 101–109 (cit. on p. 12).

- [37] Semyon Kalantarov, Raziel Riemer, and Tal Oron-Gilad. «Pedestrians' road crossing decisions and body parts' movements». In: *Transportation research part F: traffic psychology and behaviour* 53 (2018), pp. 155–171 (cit. on p. 13).
- [38] Max Theisen, Melina Bergen, Wolfgang Einhäuser, and Caroline Schießl. «Head movements predict pedestrian crossing decisions earlier than walking speed». In: *Transportation Research Part F: Traffic Psychology and Behaviour* 113 (2025), pp. 586–608 (cit. on p. 13).
- [39] Charlotte Hennah and Michail Doulas. «Dual-task walking on real-world surfaces: Adaptive changes in walking speed, step width and step height in young and older adults». In: *Experimental gerontology* 177 (2023), p. 112200 (cit. on pp. 13, 15).
- [40] Jennifer A Oxley, Elfriede Ihlen, Brian N Fildes, Judith L Charlton, and Ross H Day. «Crossing roads safely: an experimental study of age differences in gap selection by pedestrians». In: *Accident Analysis & Prevention* 37.5 (2005), pp. 962–971 (cit. on p. 13).
- [41] Aurélie Dommès and Viola Cavallo. «The role of perceptual, cognitive, and motor abilities in street-crossing decisions of young and older pedestrians». In: *Ophthalmic and physiological optics* 31.3 (2011), pp. 292–301 (cit. on p. 13).
- [42] Martina Mancini, Laurie King, Arash Salarian, Lars Holmstrom, James McNames, and Fay B Horak. «Mobility lab to assess balance and gait with synchronized body-worn sensors». In: *Journal of bioengineering & biomedical science* (2011), p. 007 (cit. on pp. 15, 16).
- [43] John Ghattas and Danielle N Jarvis. «Validity of inertial measurement units for tracking human motion: a systematic review». In: *Sports biomechanics* 23.11 (2024), pp. 1853–1866 (cit. on p. 15).
- [44] Sylvain Jung, Mona Michaud, Laurent Oudre, Eric Dorveaux, Louis Gorintin, Nicolas Vayatis, and Damien Ricard. «The use of inertial measurement units for the study of free living environment activity assessment: A literature review». In: *Sensors* 20.19 (2020), p. 5625 (cit. on p. 16).
- [45] Rachele Rossanigo, Marco Caruso, Stefano Bertuletti, Franca Deriu, Marco Knaflitz, Ugo Della Croce, and Andrea Cereatti. «Base of support, step length and stride width estimation during walking using an inertial and infrared wearable system». In: *Sensors* 23.8 (2023), p. 3921 (cit. on p. 21).
- [46] Thi-Hoa-Cuc Nguyen, Jean-Christophe Nebel, and Francisco Florez-Revuelta. «Recognition of activities of daily living with egocentric vision: A review». In: *Sensors* 16.1 (2016), p. 72 (cit. on p. 23).

- [47] Chiara Plizzari, Gabriele Goletto, Antonino Furnari, Siddhant Bansal, Francesco Ragusa, Giovanni Maria Farinella, Dima Damen, and Tatiana Tommasi. «An Outlook into the Future of Egocentric Vision: C. Plizzari et al.» In: *International Journal of Computer Vision* 132.11 (2024), pp. 4880–4936 (cit. on p. 23).
- [48] Carlo Masone and Barbara Caputo. «A survey on deep visual place recognition». In: *IEEE Access* 9 (2021), pp. 19516–19547 (cit. on p. 23).
- [49] Hamed Pirsiavash and Deva Ramanan. «Detecting activities of daily living in first-person camera views». In: *2012 IEEE conference on computer vision and pattern recognition*. IEEE. 2012, pp. 2847–2854 (cit. on p. 23).
- [50] Francesca Salis et al. «A multi-sensor wearable system for the assessment of diseased gait in real-world conditions». In: *Frontiers in Bioengineering and Biotechnology* 11 (2023), p. 1143248 (cit. on pp. 26, 27, 38, 39).
- [51] M Encarna Micó-Amigo et al. «Assessing real-world gait with digital technology? Validation, insights and recommendations from the Mobilise-D consortium». In: *Journal of neuroengineering and rehabilitation* 20.1 (2023), p. 78 (cit. on p. 26).
- [52] Cameron Kirk et al. «Mobilise-D insights to estimate real-world walking speed in multiple conditions with a wearable device». In: *Scientific reports* 14.1 (2024), p. 1754 (cit. on p. 26).
- [53] Robbin Romijnders et al. «Ecological validity of a deep learning algorithm to detect gait events from real-life walking bouts in mobility-limiting diseases». In: *Frontiers in neurology* 14 (2023), p. 1247532 (cit. on p. 26).
- [54] Paolo Tasca, Francesca Salis, Samanta Rosati, Gabriella Balestra, Claudia Mazzà, and Andrea Cereatti. «Estimating gait speed in the real world with a head-worn inertial sensor». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2025) (cit. on p. 26).
- [55] Luca Palmerini et al. «Mobility recorded by wearable devices and gold standards: the Mobilise-D procedure for data standardization». In: *Scientific Data* 10.1 (2023), p. 38 (cit. on pp. 35, 36).
- [56] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. «Places: A 10 million image database for scene recognition». In: *IEEE transactions on pattern analysis and machine intelligence* 40.6 (2017), pp. 1452–1464 (cit. on p. 40).
- [57] Glenn Jocher and Jing Qiu. *Ultralytics YOLO11*. Version 11.0.0. 2024. URL: <https://github.com/ultralytics/ultralytics> (cit. on p. 41).

- [58] Davide Ronco. «Context-Dependent Variations in Gait and Head Motion During Real-World Walking». Master's Thesis. Politecnico di Torino, July 23, 2025. URL: <https://webthesis.biblio.polito.it/36215/1/tesi.pdf> (cit. on p. 56).
- [59] Christopher L Vaughan and Mark J O'Malley. «Froude and the contribution of naval architecture to our understanding of bipedal locomotion». In: *Gait & posture* 21.3 (2005), pp. 350–362 (cit. on p. 56).
- [60] Shuen-De Wu, Chiu-Wen Wu, Shiou-Gwo Lin, Kung-Yen Lee, and Chung-Kang Peng. «Analysis of complex time series using refined composite multi-scale entropy». In: *Physics Letters A* 378.20 (2014), pp. 1369–1374 (cit. on p. 58).
- [61] Espen AF Ihlen, Aner Weiss, Alan Bourke, Jorunn L Helbostad, and Jeffrey M Hausdorff. «The complexity of daily life walking in older adult community-dwelling fallers and non-fallers». In: *Journal of biomechanics* 49.9 (2016), pp. 1420–1428 (cit. on p. 58).
- [62] Sivakumar Balasubramanian, Alejandro Melendez-Calderon, and Etienne Burdet. «A robust and sensitive metric for quantifying movement smoothness». In: *IEEE transactions on biomedical engineering* 59.8 (2011), pp. 2126–2136 (cit. on p. 59).
- [63] Paolo Brasiliano, Gaspare Pavei, and Elena Bergamini. «Smooth moves: comparing log dimensionless jerk metrics from body center of mass trajectory and wearable sensor acceleration during walking». In: *Sensors* 25.4 (2025), p. 1233 (cit. on p. 59).