



**HAL**  
open science

## A Machine Learning Pipeline for Gait Analysis in a Semi Free-Living Environment

Sylvain Jung, Nicolas de l'Escalopier, Laurent Oudre, Charles Truong, Eric Dorveaux, Louis Gorintin, Damien Ricard

► **To cite this version:**

Sylvain Jung, Nicolas de l'Escalopier, Laurent Oudre, Charles Truong, Eric Dorveaux, et al.. A Machine Learning Pipeline for Gait Analysis in a Semi Free-Living Environment. *Sensors*, 2023, 10.3390/s23084000 . hal-04362007

**HAL Id: hal-04362007**

**<https://hal.science/hal-04362007>**


Submitted on 22 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Article

# A Machine Learning Pipeline for Gait Analysis in a Semi Free-Living Environment

Sylvain Jung <sup>1,2,3,4</sup> , Nicolas de l'Escalopier <sup>5,6</sup>, Laurent Oudre <sup>1,\*</sup>, Charles Truong <sup>1</sup>, Eric Dorveaux <sup>3</sup>, Louis Gorintin <sup>7</sup> and Damien Ricard <sup>5,6,8</sup>

<sup>1</sup> Université Paris Saclay, Université Paris Cité, ENS Paris Saclay, CNRS, SSA, INSERM, Centre Borelli, F-91190 Gif-sur-Yvette, France; sylvain.jung@hotmail.com (S.J.); ctruong@ens-paris-saclay.fr (C.T.)

<sup>2</sup> Université Sorbonne Paris Nord, L2TI, UR 3043, F-93430 Villetaneuse, France

<sup>3</sup> AbilyCare, 130 Rue de Lourmel, F-75015 Paris, France; eric.dorveaux@abilycare.com

<sup>4</sup> ENGIE Lab CRIGEN, F-93249 Stains, France

<sup>5</sup> Université Paris Cité, Université Paris Saclay, ENS Paris Saclay, CNRS, SSA, INSERM, Centre Borelli, F-75006 Paris, France; ndescalopier@gmail.com (N.d.E.); damien.ricard@m4x.org (D.R.)

<sup>6</sup> Service de Neurologie, Service de Santé des Armées, HIA Percy, F-92190 Clamart, France

<sup>7</sup> Novakamp, 10-12 Avenue du Bosquet, F-95560 Baillet en France, France; lgorintin@gmail.com

<sup>8</sup> Ecole du Val-de-Grâce, Service de Santé des Armées, F-75005 Paris, France

\* Correspondence: laurent.oudre@ens-paris-saclay.fr

**Abstract:** This paper presents a novel approach to creating a graphical summary of a subject's activity during a protocol in a Semi Free-Living Environment. Thanks to this new visualization, human behavior, in particular locomotion, can now be condensed into an easy-to-read and user-friendly output. As time series collected while monitoring patients in Semi Free-Living Environments are often long and complex, our contribution relies on an innovative pipeline of signal processing methods and machine learning algorithms. Once learned, the graphical representation is able to sum up all activities present in the data and can quickly be applied to newly acquired time series. In a nutshell, raw data from inertial measurement units are first segmented into homogeneous regimes with an adaptive change-point detection procedure, then each segment is automatically labeled. Then, features are extracted from each regime, and lastly, a score is computed using these features. The final visual summary is constructed from the scores of the activities and their comparisons to healthy models. This graphical output is a detailed, adaptive, and structured visualization that helps better understand the salient events in a complex gait protocol.

**Keywords:** free-living; wearable sensor; IMU; graphical feedback; change point detection; Human Activity Recognition



**Citation:** Jung, S.; de l'Escalopier, N.; Oudre, L.; Truong, C.; Dorveaux, E.; Gorintin, L.; Ricard, D. A Machine Learning Pipeline for Gait Analysis in a Semi Free-Living Environment. *Sensors* **2023**, *1*, 0. <https://doi.org/>

Academic Editor: Arnold Baca

Received: 8 March 2023

Revised: 3 April 2023

Accepted: 13 April 2023

Published:



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Portable inertial sensors such as accelerometers, gyroscopes, or Inertial Measurement Units (IMUs) are frequently used to analyze human activity. In particular, gait quantification is a major subject of interest for clinicians, as it can help to precociously detect the risk of falling or be applied in the context of longitudinal follow-up for degenerative diseases [1]. Most of the published studies are led in clinical and controlled environments (laboratories, etc.), where efficient algorithms have been developed to extract from the raw data relevant features such as gait events, with precision as low as a few milliseconds [2–6].

However, these environments can induce a Hawthorne effect (“white coat” effect) on the patients and therefore bias the gait analysis process [7,8]. Indeed, during controlled studies, subjects are aware of being measured and this implies over- or under-performance during the protocol. A study has shown for instance that the variability of step duration in a free or semi-controlled environment is statistically different from the one measured in a controlled environment [9]. These effects can be due to the bulkiness of the sensors,

to the instructions given to the recorded subjects, or to the narrowness of the measurement environments. Moreover, gait cannot be fully apprehended and analyzed without acknowledging phenomena caused by free-living measurements (fatigue, open spaces). A wide range of new protocols, referred to as Free-Living Environments (FLEs) protocols have therefore emerged to avoid these issues and to improve our understanding of human behavior in FLEs, or in the "wild" [10].

The study of locomotion is more challenging in FLEs than in controlled settings, as the automatic computation of gait features at a micro level in FLE (stride times, step times) is more complex when no context information is available [11]. As for the current studies set up by the scientific community in FLEs, they present very diversified goals. The vast majority of studies carried out in FLEs aim to perform Human Activity Recognition (HAR) and to use its outputs to enable clinical analyses [12,13]. HAR can thus help providing general metrics about activity durations in order to enable a quantified follow-up of patients' physical activity. This can also be the first step in estimating a patient's energy expenditure over long periods. Indeed, some works aim specifically to measure the time spent in activities that require a greater or lesser expenditure of energy with or without using HAR [14–16]. Activities are then associated with a metabolic equivalent of task which is the objective measure of a participant's rate of energy expenditure in relation to mass. Finally, another large group of studies is dedicated to the detection of falls in FLEs in order to respond to the public health issues that these falls represent [17,18].

In almost all of the work cited above, the output measures provided to clinicians are often generalized/aggregated through simple features. For instance, HAR studies tend to only determine the time spent in various targeted activities, whereas studies that focus on calculating energy expenditure only quantify the time spent in more or less energy-consuming activities. Works that seek to detect falls rely on general output metrics as well (number of falls, time of falls). This, therefore, leads to analyses of physical activity summarized by aggregate and averaged output measures that may hide some relevant phenomena of interest. To get around this drawback, one idea would be to use fine-grained features such as those used in clinical settings. However, this would imply a greater computational burden (e.g., detection of all steps, strides) and a prohibitive flow of information that would overwhelm clinicians and prevent them from obtaining a clear and quickly understandable assessment of the physical activity of their patients. Based on this observation, we present in this paper an alternative and intermediate solution whose major innovation is to provide an accurate macro-analysis with a low computational cost and which is ergonomic for clinicians. This visual summary also allows to keep a temporal structure that helps to provide new interpretations to the free-living recordings (impact of fatigue on specific regimes, performance during transitions between long and short walking regimes).

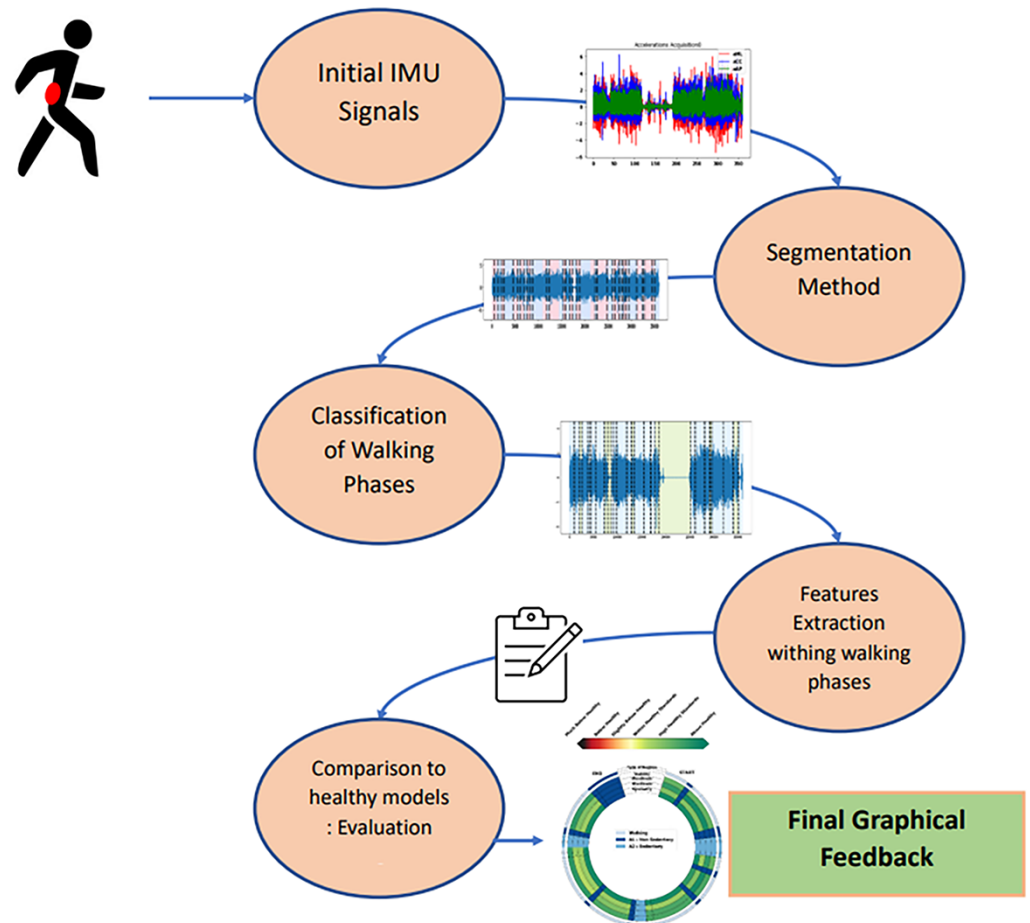
To the authors' knowledge, few, if any studies based on the use of IMUs in FLEs have endeavored to provide a macro-analysis displayed in the form of an easy-to-understand visual summary that fully assesses the entire timeline of FLE signals. The purpose of our paper is to describe the processing steps to compute such accurate and didactic quantitative feedback on a subject's walking during an acquisition in a Semi Free-Living Environment (Semi FLE). We believe that such visualization will help clinicians to perform more accurate and comprehensive longitudinal tracking of locomotion in the natural environment of their patients. This could for instance allow us to evaluate rehabilitation procedures or treatment choices for specific diseases and to assess the impact of treatments on pathologies such as musculoskeletal tumors of the lower limbs or neurological disorders (Parkinson's disease for instance). The current monitoring of the effects of these FLE treatments is only carried out via calculations of general metrics (steps/day, ambulatory bouts/day) [19] when our graphical tool will enable a refined follow-up displaying an enhanced macro analysis of gait phases.

This paper is organized as follows. Section 2 describes the protocols and algorithms used to create the graphical tool in question. Section 3 provides a comprehensive evaluation

of each of the steps of the processing pipeline as well as some results obtained on both pathological and healthy subjects. In Section 4, these results are analyzed and discussed.

## 2. Material and Methods

The proposed pipeline is composed of four successive steps: data segmentation, data classification, feature extraction, and comparison with a healthy model. Figure 1 summarizes these successive steps from the raw data to the final graphical feedback.



**Figure 1.** Successive steps of our processing pipeline to render a graphical feedback from a semi-FLE acquisition.

- The segmentation step uses an adaptive change-point detection algorithm to process IMU recordings. The method searches for significant changes in the time-frequency space at a given scale, i.e., instants where the subject changed their behavior/activity. Signals are thus segmented into several homogeneous regimes that will help to extract knowledge from the global recording.
- Once these homogeneous regimes are segmented, they are classified as walking or non-walking phases through a supervised classification procedure. A second algorithm identifies, within non walking phases, sedentary and non-sedentary regimes, thus providing a full labelization of the regimes. Sedentary regimes correspond to activities that are not walking phases but that imply movements from the recorded subjects (in our case, walking up and down stairs, opening a fire door, and performing a 90° turn). On the other hand, non-sedentary regimes correspond to activities that do not imply movements from participants (in our case, leaning, sitting, standing).
- The next step consists of extracting features from the regimes that have been classified as corresponding to a walking phase. These features were selected in order to assess different aspects of gait (stability, steadiness, sturdiness, and symmetry).

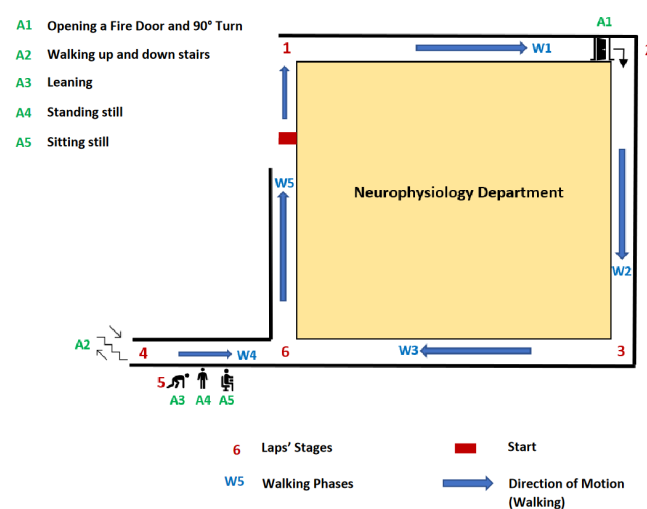
- By using models learned from healthy subjects, each walking regime is then given a score represented by distinct color, allowing visual and intuitive feedback.

### 2.1. Data and Protocols

A total of 21 healthy subjects ( $33.4 \pm 14.42$  years old, 10 men and 11 women), 6 patients having undergone or about to undergo an orthopedical surgery ( $48.75 \pm 20.32$  years old, 1 man and 5 women), and 3 patients suffering from a neurological pathology due to a cerebral lesion ( $71.5 \pm 8.22$  years old, 1 man and 2 women) were measured on a semi-controlled protocol. Informed consent was obtained for all participants.

Subjects were equipped with a Shimmer3 IMU (Shimmer, Dublin, Ireland) on their lower back L5 (via the use of a belt strap) (with sampling frequency  $F_s = 100$  Hz, battery life = 39–69 h, storage = 8 GB,  $\pm 2$  g (to  $\pm 16$  g) for the accelerometer,  $\pm 250$  dps (to  $\pm 2000$  dps) for the gyroscope [10]. Subjects were asked to complete several laps of the Neurophysiology Department at Percy Hospital (a semi-controlled environment) and to perform activities at the end of each lap (climbing up and down some stairs, leaning, sitting, standing) for a total protocol duration of precisely 6 min. Participants can complete different numbers of laps depending on their pathology or walking characteristics. However, subjects must be able to complete at least one full lap in less than 6 min in order to perform at least one end-of-lap sedentary activity. If a registered subject fails to complete a lap of the protocol, he/she is not included in the list of selected subjects. Raw signals are filtered between 0.5 and 5 Hz to remove the noise [20] with a Butterworth bandpass filter (4th order). This experimental protocol was approved by the committee for the protection of individuals (Comité de Protection des Personnes) from the Agence Régionale de Santé (ARS). The ID-RCB number of the committee in which this study is included is 2021-A00087-34. All methods were carried out in accordance with the principles of the Declaration of Helsinki.

The protocol, illustrated on Figure 2, contains several regimes that are either walking phases (denoted  $W\bullet$ ) or activities (denoted  $A\bullet$ ). During each recording, breakpoints were identified thanks to a camera carried by the subject who was left alone to perform the protocol. These breakpoints correspond to transitions between walking phases/activities, activities/walking phases, activities/activities, or walking phases/walking phases. These annotations were conducted collaboratively by two experts and consist of precise timestamps that will be used as ground truth breakpoints' labels. Characteristics of these transitions are displayed in Table 1.



**Figure 2.** Description of the semi-controlled protocol. Numbers displayed indicate the position of the subject during its path.

**Table 1.** Details of annotated transitions.

Transition Identification	Details	Type of Regime
W1	Walk (1 → 2)	Walking
A1	Door opening and 90-degree turn	Non-Sedentary
W2	Walk (2 → 3)	Walking
W3	Walk (3 → 4)	Walking
A2	Going up 3 stairs U-turn and going down 3 steps stairs	Non-Sedentary
A3	Leaning	Sedentary
A4	Standing Still	Sedentary
A5	Sitting Still	Sedentary
W4	Walk (5 → 6)	Walking
W5	Walk (6 → 1)	Walking

## 2.2. Step 1: Adaptive Changepoint Detection Method

The first step of the processing pipeline consists of segmenting the raw signals by detecting all activity changes. Intuitively, this segmentation depends on the meaning given to the notion of *change*. In order to adapt the strategy to the signals of interest, we propose to use a supervised approach that learns from annotated data the type of changes that are meaningful in the context and thus the adequate *scale* for the segmentation algorithm.

### 2.2.1. Data Transformation

In the proposed method, changes are detected in the short-time Fourier transform (STFT) of the raw acceleration/angular velocity signals recorded on the lower back of healthy subjects. This data transformation has been successfully used in multiple publications as it enables visualizing activity changes but also speed changes or event slope changes [21–23].

More specifically, two signals of interest are extracted from the raw data (IMU recordings): the craniocaudal angular velocity (gCC) and the anteroposterior acceleration (aAP). These signals were chosen as they are directly influenced by the changes that can be observed during the execution of the protocol (beginning and end of gait, short activities, half turns, and turns). Studies have already used these signals to meet objectives similar to ours (turn detection, detection of activities of daily life) [24,25]. They are normalized (zero mean and unit variance) before being transformed in the time-frequency domain through STFT (3 s window length and 0.1 hop length). Only the 0–5 Hz frequency band, where phenomena of interest are contained, is kept [26,27]. The norms of the STFT coefficients of each aAP and gCC signal are computed and concatenated, providing  $d = 28$  frequency bins per frame (14 per signal). The output data will be seen by the segmentation algorithm as a  $d$ -dimensional multivariate signal.

### 2.2.2. Changepoint Detection Algorithm

Formally, let  $y = \{y_1, y_2, \dots, y_n\}$  denote a  $\mathbb{R}^d$ -valued signal with  $n$  frames: the goal is to detect shifts in the mean of this signal. For  $K$  change point indexes  $\{t_k\}_{k=1}^K$  ( $1 < t_1 < t_2 < \dots < t_K < n$ ), a common measure of approximation quality is the empirical quadratic risk:

$$R(y, \{t_k\}) := \sum_{k=0}^{K+1} \left( \sum_{t=t_k}^{t_{k+1}-1} \|y_t - \bar{y}_{t_k..t_{k+1}}\|^2 \right) \quad (1)$$

where  $\bar{y}_{t_k..t_{k+1}}$  is the empirical mean of  $y_{t_k}, \dots, y_{t_{k+1}-1}$  and  $t_0 := 1$  and  $t_{K+1} := n$  by convention. The risk (1) is simply the error when approximating  $y$  by a piecewise constant

signal. The objective is to find the change points  $\{t_k\}_{k=1}^K$  that minimize this risk. When the number of breaks  $K$  is not known (which is the case in this article), the empirical quadratic risk is penalized with a linear penalty and the optimal breakpoints can be estimated by solving the following optimization problem:

$$\hat{K}, \hat{t}_1, \dots, \hat{t}_K := \arg \min_{K, t_1, \dots, t_K} \left( \underbrace{R(y, \{t_k\}_{k=1}^K)}_{R_\beta(y, \{t_k\})} + \beta K \right) \quad (2)$$

where the smoothing parameter  $\beta > 0$  controls the trade-off between model complexity and model accuracy. The value of  $\beta$  is critical: a large  $\beta$  only detects strong breaks (change of activity for instance) whereas a low  $\beta$  also detects small breaks (change within the walking phases). For a fixed  $\beta$ , this discrete optimization problem can be solved efficiently in linear time  $\mathcal{O}(n)$  with the PELT algorithm [28].

### 2.2.3. Calibration of $\beta$

Instead of manually calibrating this parameter by trial and error, a supervised approach, described in previous studies [29,30], is adopted in this article to learn an optimal value. This method takes as input an annotated database of physiological signals where the relevant breakpoints have been annotated by an expert, and learns the adequate parameter value that is able to reproduce the segmentation strategy of the expert on new signals.

Formally, the input is a collection of  $N$  labeled signals  $y^{(1)}, \dots, y^{(N)}$ , and for each  $y^{(i)}$ , an expert manually provided the set of true change point indexes  $\{t_k^{(i)}\}_k$ . The optimal smoothing parameter, denoted  $\hat{\beta}_{\text{opt}}$ , is such that the risk of the true expert segmentation on all signals  $R_\beta(y^{(i)}, \{t_k^{(i)}\}_k)$  is closest to the one of the predicted segmentation on all signals  $R_\beta(y^{(i)}, \{\hat{t}_k^{(i)}\}_k)$ :

$$\hat{\beta}_{\text{opt}} := \arg \min_{\beta > 0} \frac{1}{N} \sum_{i=1}^N \left( R_\beta(y^{(i)}, \{t_k^{(i)}\}_k) - R_\beta(y^{(i)}, \{\hat{t}_k^{(i)}\}_k) \right). \quad (3)$$

Intuitively, the algorithm will search for the penalty  $\beta$  that allows to reproduce the annotations by forcing the  $\beta$ -optimal solution  $\{\hat{t}_k^{(i)}\}_k$  to be as close as possible to the ground truth partition  $\{t_k^{(i)}\}_k$ . The excess penalized risk is a convex function w.r.t.  $\beta$  (precisely, this is an affine function minus a concave function: see [29] for details). We therefore used Brent's method as a convex optimization tool to minimize this component for each signal.

### 2.3. Step 2: Classification of Segmented Phases

Once the homogeneous regimes have been extracted from the raw data, our aim is to assess whether each of these regimes consists of a walking phase or another activity (sedentary or not). This task is known in the literature as Walking Bout (WB) detection.

There are several existing approaches to perform this detection. The first two approaches are peak detection [31] and step detection methods [32]. These two types of methods impose a significant computational heaviness and are inadequate for FLEs if we wish to provide a simple and easy-to-understand visualization method as we do. Moreover, these types of studies do not allow to characterize all the portions of the signal (only the WBs) as it is desired in our case. A third group of studies integrates the detection of WB into HAR methods via the use of classifiers. To do so, a list of features is extracted on filtered signals in each frame selected by sliding windows. Machine learning classifiers are trained with supervised data (features' values associated with specifically targeted labels) and then implemented for each portion of the signal. Several classifiers in particular are then used with a very satisfactory accuracy rate: the support vector machine (SVM) [33], random forest [34], decision trees [35] and k-nearest neighbors [36]. This way of detecting

WBs using classifiers matches the objectives of our study. Indeed, it will be possible by using this approach to label every portion of the signal and not only the walking regimes (unlike the first two approaches). Moreover, the computational costs are lower than the first two approaches with satisfactory results.

Our classification procedure is therefore based on feature extraction coupled with supervised learning. However, contrary to state-of-the-art methods used for HAR, instead of using (possibly overlapping) frames, we propose to perform this classification at the regime level. The advantages are twofold: first, because of the segmentation procedure, we know that each regime is stationary, which is a valuable theoretical property for computing robust features. Second, the average length of the regimes is often longer than typical frame durations, which provides more data for computing features.

First, we extract for each detected regime an extensive list of both temporal and frequency features such as variances, means, dominant frequencies, and power at dominant frequencies. These features have been chosen and selected in accordance with a recent state-of-the-art article [10] dedicated to activity classification from IMU signals. The features list is presented in Table 2. These 135 features are retained because their computation is convenient for long FLE signals as they do not require any detection of events (heel strikes, toe strikes). PCA is then applied to these 135 features in order to keep 99% of the cumulative explained variance. As for classifiers, we have compared their performance when applied to our signals, and the SVM classifier with linear kernel was retained for our study.

**Table 2.** Details of features used for classification. For each formula,  $X = [x_1, x_2 \dots, x_n]$  is assumed to be one of the six dimensions (three linear accelerations and three angular velocities) from the IMU signal. In total, 135 features are used. Notations:  $\bar{x}$  is the empirical mean of  $X$ ,  $\hat{\sigma}$  is the empirical unbiased standard deviation of  $X$ ,  $\text{FFT}(X)$  is the Fourier transform of  $X$ ,  $\text{ConjFFT}(X)$  denotes the complex conjugate of  $\text{FFT}(X)$ .

Features	Signal Used for Computation	Description	Domain	Formulas
mean_signal	All 6 Signals	Mean	Time	$\bar{x}$
std_signal	All 6 Signals	Standard deviation	Time	$\hat{\sigma}$
var_signal	All 6 Signals	Variance	Time	$\hat{\sigma}^2$
min_signal	All 6 Signals	Minimum	Time	$\min(x)$
max_signal	All 6 Signals	Maximum	Time	$\max(x)$
PD0_signal	All 6 Signals	Power at the first dominant frequency	Frequency	$\max\left(\frac{\text{FFT}(X) * \text{ConjFFT}(X)}{N}\right)$
F0_signal	All 6 Signals	First dominant frequency	Frequency	$\text{argmax}\left(\frac{\text{FFT}(X) * \text{ConjFFT}(X)}{N}\right)$
PD2_signal	All 6 Signals	Power at the second dominant frequency	Frequency	$\text{second\_max}\left(\frac{\text{FFT}(X) * \text{ConjFFT}(X)}{N}\right)$
F2_signal	All 6 Signals	Second dominant frequency	Frequency	$\text{second\_argmax}\left(\frac{\text{FFT}(X) * \text{ConjFFT}(X)}{N}\right)$
CV_signal	All 6 Signals	Coefficient of variation	Time	$\hat{\sigma} / \bar{x}$



Table 2. Cont.

Features	Signal Used for Computation	Description	Domain	Formulas
p75_signal	All 6 Signals	75th percentile	Time	Let R be the 75th percentile rank: $R = \frac{75*n}{100}$ , p75 corresponds to the Rth value on the sorted X array
p25_signal	All 6 Signals	25th percentile	Time	Let R be the 25th percentile rank: $R = \frac{25*n}{100}$ , p25 corresponds to the Rth value on the sorted X array
p85_signal	All 6 Signals	85th percentile	Time	Let R be the 85th percentile rank: $R = \frac{85*n}{100}$ , p85 corresponds to the Rth value on the sorted X array
p15_signal	All 6 Signals	15th percentile	Time	Let R be the 15th percentile rank: $R = \frac{15*n}{100}$ , p15 corresponds to the Rth value on the sorted X array
p95_signal	All 6 Signals	95th percentile	Time	Let R be the 95th percentile rank: $R = \frac{95*n}{100}$ , p95 corresponds to the Rth value on the sorted X array
p5_signal	All 6 Signals	5th percentile	Time	Let R be the 5th percentile rank: $R = \frac{5*n}{100}$ , p5 corresponds to the Rth value on the sorted X array
p75m_signal	All 6 Signals	75th percentile at the middle of the signal (2/3 of the signal)	Time	Let R be the 75th percentile rank: $R = \frac{75*n}{100}$ , p75 corresponds to the Rth value on the sorted X array
p25m_signal	All 6 Signals	25th percentile at the middle of the signal	Time	Let R be the 25th percentile rank: $R = \frac{25*n}{100}$ , p25 corresponds to the Rth value on the sorted X array
p85m_signal	All 6 Signals	85th percentile at the middle of the signal	Time	Let R be the 85th percentile rank: $R = \frac{85*n}{100}$ , p85 corresponds to the Rth value on the sorted X array
p15m_signal	All 6 Signals	15th percentile at the middle of the signal	Time	Let R be the 15th percentile rank: $R = \frac{15*n}{100}$ , p15 corresponds to the Rth value on the sorted X array
RMS_signal	All 6 Signals	Root mean Ssquare	Time	$RMS = \sqrt{\frac{1}{n} * (\sum_{i=1}^n x_i^2)}$
P1_aCC	aCC	First peak of autocorrelation coefficients for craniocaudal acceleration	Time	$ACF = iFFT[FFT(X) * ConjFFT(X)]$ , P1 is the first peak of ACF
P2_aCC	aCC	Second peak of autocorrelation coefficients for craniocaudal acceleration	Time	$ACF = iFFT[FFT(X) * ConjFFT(X)]$ , P2 is the second peak of ACF
VM	All 6 Signals	Vector magnitude of all accelerations (craniocaudal aCC, mediolateral aML, and anteroposterior aAP)	Time	$VM = \sqrt{aCC^2 + aAP^2 + aML^2}$

In our pipeline, two classifiers are trained and used in cascade. The first SVM classifier with the linear kernel (Classifier 1) performs a walking/non-walking binary classification,

whereas the second SVM classifier with linear kernel (Classifier 2) only considers the non-walking phases detected by Classifier 1 and classifies them as sedentary or non sedentary.

#### 2.4. Step 3: Feature Extraction

In our visual summary, a walking regime is assessed according to four standard criteria: stability, steadiness, sturdiness, and symmetry. The four features to be extracted for regime evaluation, detailed in Table 3, are chosen because of their ability to accurately characterize the gait, their recurrent use in the literature, and ease of computation. These features are directly calculated on filtered signals at each regime level before the PCA has been applied.

**Stability** [37,38]: criterion evaluating postural balance and used for instance to prevent falls [39]. Stable walking can be defined as *gait that does not lead to falls despite perturbations* [40]. This aspect is evaluated by using the root mean square ratio computed on the mediolateral acceleration  $RMSR_{ML}$ . It corresponds to the ratio of the root mean square of the mediolateral accelerations  $RMS_{ML}$  to the root mean square vector magnitude computed on all axes  $RMS_A$  as displayed in (4). RMS evaluates the magnitude of the acceleration on one specific axis. The higher  $RMSR_{ML}$  is, the higher the values of mediolateral accelerations tend to be compared to other accelerations. This indicates an instability on the mediolateral axis and therefore a postural instability. Thus,  $RMSR_{ML}$  is selected for our study since it has been proven to be uncorrelated with walking speed [38].

$$\begin{aligned} RMS_A &= \sqrt{RMS_{ML}^2 + RMS_{CC}^2 + RMS_{AP}^2}, \\ RMSR_{ML} &= \frac{RMS_{ML}}{RMS_A} \end{aligned} \quad (4)$$

**Sturdiness** [3,41]: criterion *evaluating gait amplitude* [3]. For instance, sturdiness can be assessed to quantify observed defects in patients with Parkinson's disease with low amplitude movements [42]. This aspect is evaluated by using the root mean square ratio computed on the anteroposterior acceleration  $RMSR_{AP}$  whose computation is conducted the same way it is performed for  $RMSR_{ML}$ . The higher it is, the higher the anteroposterior accelerations' values compared to other accelerations and the higher the sturdiness is. Indeed, high anteroposterior acceleration values mean that step impulses are vigorously performed by the participant.  $RMSR_{AP}$  is used instead of  $RMS_{AP}$  in order to limit the influence of the walking speed.

$$RMSR_{AP} = \frac{RMS_{AP}}{RMS_A} \quad (5)$$

**Steadiness**[43,44]: criterion evaluating step regularity, i.e., *similarity of consecutive strides*[45]. This criterion can be used to quantify locomotion flaws in targeted cohorts with lower limb defects (such as transfemoral amputees). This category is evaluated by using the second peak of the autocorrelation coefficients ( $P2_{CC}$ ) calculated on craniocaudal accelerations via the Wiener–Khinchin theorem. This unbiased autocorrelation function uses both fast Fourier transform (FFT) and inverse fast Fourier transform (iFFT) as detailed in Table 3. This feature compares the similarity between strides within a walking regime since it occurs with a time lag of two steps. The higher  $P2_{CC}$  is, the more similar the performed strides are. Let  $aCC$  be the associated craniocaudal acceleration signal,  $ConjFFT(aCC)$  the complex conjugate of  $FFT(aCC)$ , and  $ACF$  the autocorrelation coefficients:  $P2_{CC}$  is defined as detailed in (6). Figure 3 shows a craniocaudal acceleration signal associated with its autocorrelation:  $P1_{CC}$  and  $P2_{CC}$  locations are presented.

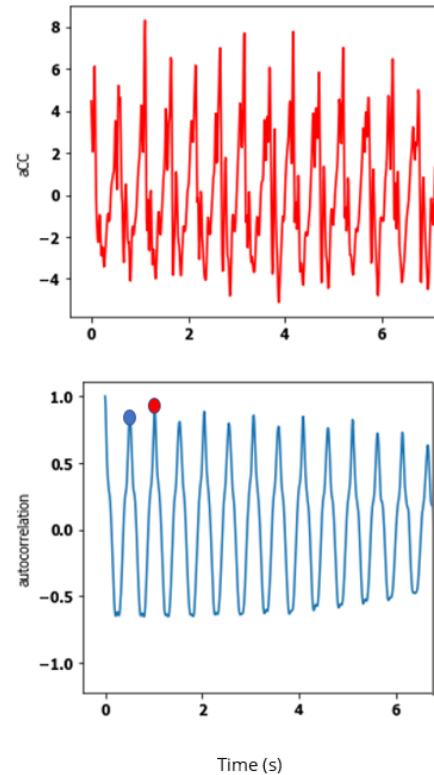
$$ACF = iFFT[FFT(aCC)ConjFFT(aCC)] \quad (6)$$

**Symmetry**[43,44]: criterion evaluating step symmetry. *A symmetric gait pattern for humans is characterized by the almost identical behavior of bilateral limbs during a gait cycle* [46]. This aspect

is evaluated by using the first peak of the autocorrelation coefficients ( $P1_{CC}$ ) calculated on craniocaudal accelerations via the Wiener–Khinchin theorem. It compares the similarity between steps within a walking regime since it occurs with a time lag of one step.  $P1_{CC}$  evaluates the ability to maintain vertical correspondence between right and left hemi-bodies during walking regimes. The higher it is, the more similar the steps from both sides are.

**Table 3.** Features used to establish scores for the graphical feedback. In total, four features are used.  $\text{ConjFFT}(X)$  denotes the complex conjugate of  $\text{FFT}(X)$ .

Categories	Features	Description	Mathematical Computation
Steadiness	$P2_{CC}$	The second peak of the autocorrelation coefficients calculated on craniocaudal accelerations via the Wiener–Khinchin theorem: the higher it is, the more similar the steps are.	$ACF = iFFT[FFT(X)ConjFFT(X)]$ , P1 is the first peak of ACF whereas P2 is the second peak
Symmetry	$P1_{CC}$	The first peak of the autocorrelation coefficients calculated on craniocaudal accelerations via the Wiener–Khinchin theorem: the higher it is, the more similar the strides are.	P1 is the first peak of ACF whereas P2 is the second peak
Sturdiness	$RMSR_{AP}$	Root mean square ratio on anteroposterior acceleration. The higher it is, the higher the sturdiness is.	$RMS_A = \sqrt{RMS_{ML}^2 + RMS_{CC}^2 + RMS_{AP}^2}$ $RMSR_{AP} = \frac{RMS_{AP}}{RMS_A}$
Stability	$RMSR_{ML}$	Root mean square ratio on mediolateral acceleration. The lower it is, the higher the stability is.	$RMS_A = \sqrt{RMS_{ML}^2 + RMS_{CC}^2 + RMS_{AP}^2}$ $RMSR_{ML} = \frac{RMS_{ML}}{RMS_A}$



**Figure 3.** One aCC signal and its associated unbiased autocorrelation. Definition of  $P1_{CC}$  (blue dot) and  $P2_{CC}$  features (red dot).

### 2.5. Step 4: Score Generation and Graphical Feedback

Using a database of healthy walking phases taken from the healthy subjects from the recorded cohort, statistics for the different features are computed (means, percentiles, etc.). These models are then used to assess each novel walking phase with the scoring procedure described as follows.

Considering a feature with mean  $\mu$  and standard deviation  $\sigma$  on all walking regimes from healthy subjects, we compute the z-score normalized feature

$$z = \frac{x - \mu}{\sigma}. \quad (7)$$

The z-score normalized features are then displayed with a color bar of boundaries  $[x_{min}, x_{max}]$  where  $x_{min}$  and  $x_{max}$  are, respectively, the 10% and 90% percentiles of the normalized features on the healthy subjects. *Slightly below healthy* values correspond to regimes with z-scores just below those obtained for the 10th percentile of healthy subjects, and *slightly above healthy* values correspond to regimes with z-scores just above those obtained for the 90th percentile of healthy subjects.

### 2.6. Evaluation Metrics

In the next section, we conduct a step-by-step assessment of each step of the pipeline with adapted metrics. All simulations are run with 3-fold cross-validation. To that aim, we split the dataset into three balanced sets (two training sets and one validation set) of seven healthy subjects and three pathological subjects (two participants having undergone or about to undergo orthopedical surgery and one neurological patient). This cross-validation allows one to verify that the algorithm developed in this study can be used on new unseen data to apply the desired segmentation.

#### 2.6.1. Evaluation of the Adaptive Change-Point Detection

The supervised segmentation procedure is assessed with three standard evaluation metrics: precision, recall, and F1-score. A predicted change point is a true positive (TP) if it is close to a true change point (within a specific positive temporal margin). This margin is set to 3.5 s and corresponds to the maximum accepted error for a change point. It must be lesser or equal to the minimum temporal distance between two true change points. Recall which corresponds to the proportion of true change points that are correctly predicted is the ratio of the number of TPs to the number of true change points  $K^*$ . Precision is the proportion of predicted change points that are associated with true change points. It is the ratio of the number of TPs to the number of predicted change points  $K$ .

$$Precision = \frac{TP}{K} \quad Recall = \frac{TP}{K^*} \quad F1Score = 2 \frac{Precision Recall}{Precision + Recall} \quad (8)$$

#### 2.6.2. Joint Evaluation of Segmentation and Classification Steps

After the first two steps of the pipeline, each data sample is labeled as Walking, Non-Walking/Non-Sedentary, or Sedentary. Intuitively, the sample-scale classification performances depend both on the segmentation step and on the regime classification step. To jointly evaluate these tasks, we compute the confusion matrix between all three labels. Each coefficient of the matrix represents the percentages of samples annotated as belonging to the row activity that have been classified as the column activity. Perfect performances would correspond to a diagonal matrix.

## 3. Results

### 3.1. Adaptive Change-Point Detection

The cross-validation results are the following: **F1-Score**  $0.76 \pm 0.01$ , **Recall**  $0.79 \pm 0.02$ , and **Precision**  $0.74 \pm 0.02$ . These are satisfactory results and enable a proper segmentation that can be relied on for the further setup of graphical feedback. The recall

and precision results are well-balanced, which means that there is no oversegmentation or undersegmentation.

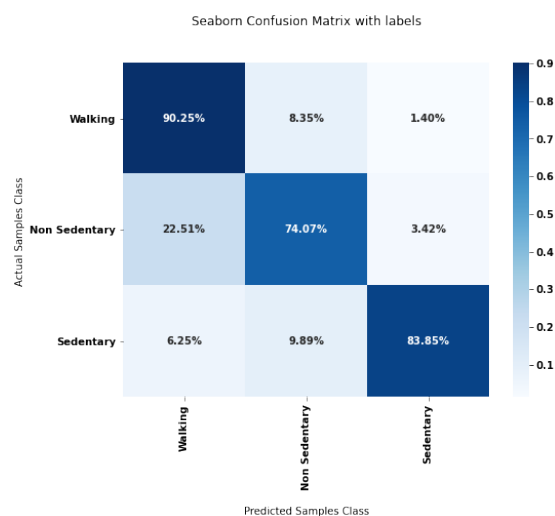
### 3.2. Joint Evaluation of Segmentation and Classification Steps

Figure 4 displays the confusion matrix at sample scale evaluating both segmentation and classification. The correct class (Walking, Non-Sedentary, Sedentary) is accurately predicted for a large majority of samples. In detail, Walking samples are well distinguished from other classes whereas Sedentary and Non-Sedentary samples are less precisely discriminated.

Several HAR methods detailed in the literature attempt to classify WBs. Their results have all been evaluated in a review conducted by the authors of the current article [10]. Table 4 displays accuracy values for different kinds of classifiers used in recent HAR studies carried out in free-living conditions as well for these same classifiers tested on our retrieved data. Accuracy results are satisfactory for our implemented chosen linear SVM cascade classifiers (0.88) when compared to these results (no other accuracy result above 0.89). It is noteworthy that the majority of these methods are difficult to compare with ours because of variations in the conditions of application of the classifiers (very few studies compute the training features at the scale of a regime for instance, sensors are placed on different locations).

**Table 4.** Average accuracy values for different kinds of classifiers used in the literature to classify activities (including walking bouts) in free environments.

Type of Classifiers	Reported Performances	Performances on Our Data
Support Vector Machine SVM	0.72 [47]	0.88 ± 0.14
	0.85 [14]	
	0.74 [48]	
Random Forest	0.88 [49]	0.85 ± 0.16
	0.88 [50]	
	0.86 [14]	
Decision Tree	0.82 [47]	0.77 ± 0.17
	0.83 [51]	
	0.80 [14]	
k Nearest Neighbors	0.75 [47]	0.89 ± 0.06
	0.74 [49]	
	0.68 [50]	



**Figure 4.** Confusion matrix for the segmentation and classification steps of our processing pipeline.

### 3.3. Scores and Graphical Feedback

Graphical outputs are plotted for four subjects: one 23 year-old healthy male subject (*HSU*—Figure 5a), one 66 year-old female pathological subject (*PSU1*—Figure 5b) with a chronic gluteus medius insufficiency, one 75 year-old woman with post-radiation left brachial plexitis called *PSU2*, and one 25 year-old female pathological subject (*PSU3*—Figure 6b) in an immediate post-operative phase of a knee ligamentoplasty. For the *PSU3* subject, we have also computed the visual feedback in the immediate pre-operative phase of her knee ligamentoplasty as displayed in Figure 6.

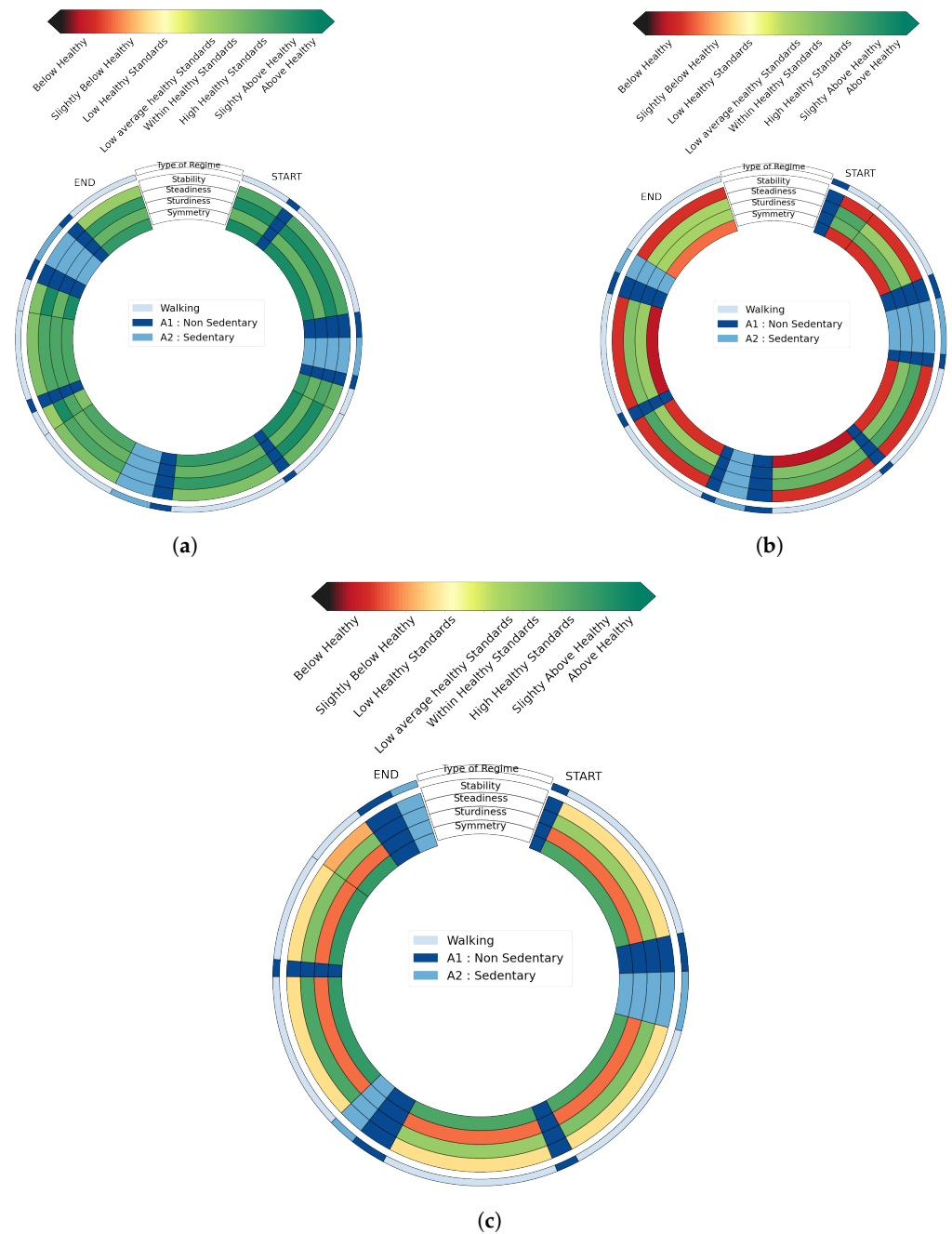
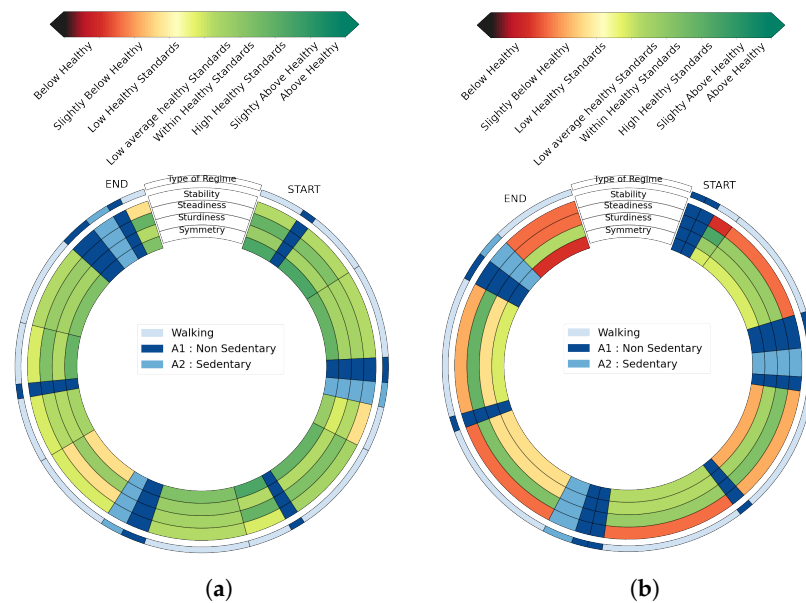


Figure 5. Graphical feedbacks from (a) *HSU*, (b) *PSU1*, (c) *PSU2*.



**Figure 6.** Evaluation of pre-hospitalization acquisition and post-hospitalization acquisition for a subject who has undergone knee ligamentoplasty. Post-operation evaluation displays a worse state. (a) Pathological Subject 3 graphical feedback **Pre-surgery** PSU3A; (b) Pathological Subject 3 graphical feedback: **Post-surgery** PSU3B.

These graphs display the evaluation of the whole protocol segmented in regimes in a clockwise manner. The first outer circle specifies the nature of the segmented regimes: dark blue for non-sedentary activities, standard blue for sedentary activities, and light blue for walking regimes. The four next inner concentric circles are each associated with a gait criterion: stability, steadiness, sturdiness, and symmetry. Each portion of these circles delimited by black lines corresponds to a segmented regime, whose length is proportional to the duration of the regime. For a given evaluation criterion, each walking regime is then assigned a color from dark red to dark green. This color depends on the comparison of this regime to the average healthy walking regime. Non-walking regimes are not evaluated and are displayed in dark blue for non-sedentary activities and light blue for sedentary activities (as in the first outer circle).

Several interesting features can be highlighted by simple visual inspections of these diagrams. We see for instance in the feedback obtained for *HSU* (Figure 5a) that walking regimes from this participant are deemed to be of satisfactory quality according to all evaluated criteria. On the other hand, scores' figures for several walking regimes from *PSU1* and *PSU2* are below average healthy standards. For *PSU1*, stability and symmetry (first and fourth inner circles) are degraded whereas *PSU2* displays deteriorated stability and sturdiness (first and third inner circles). The tool thus allows for a simple interindividual comparison, which is both quantitative (thanks to the colors) and qualitative (thanks to the four criteria). Furthermore, the temporality of the whole exercise is preserved, allowing for a better understanding and interpretation. Figure 6 shows another potential use for the tool in the context of longitudinal follow-up. By comparing the diagram obtained before (Figure 6a) and after surgery (Figure 6b) from participant *PSU3*, it is visible that the post-surgery consequences mostly affect the stability of the locomotion, and that symmetry is also degraded.

## 4. Discussion

### 4.1. Performances

The metrics obtained for the supervised segmentation are satisfactory (F1-score around 75%) and the observed margin of error is due either to annotation approximations or specific breakpoints difficult to detect because of the movement of the subjects (quick transitions or turns that are taken so quickly that they do not appear as a clear breakpoint). A study

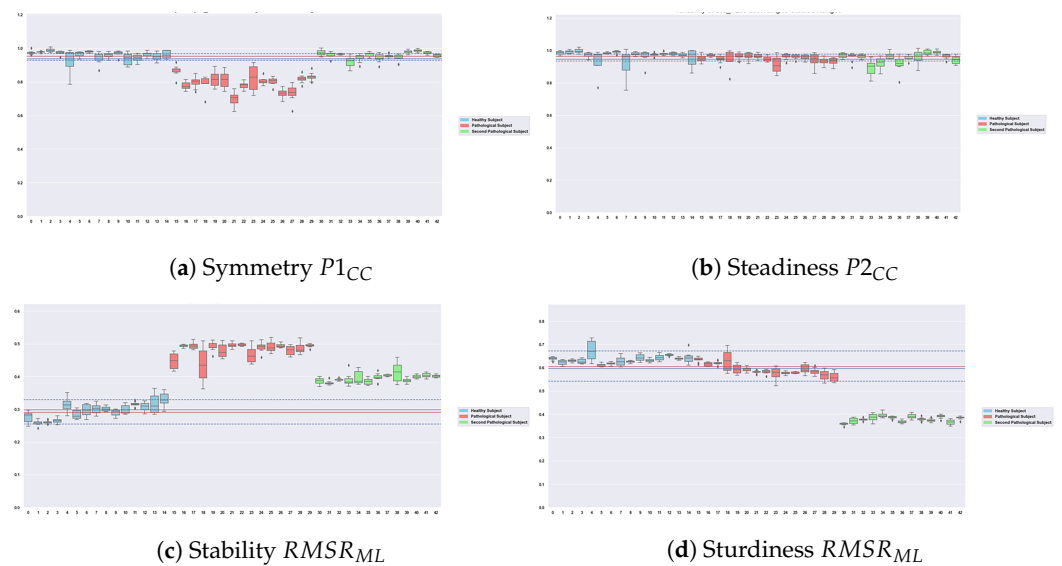
differentiating the results obtained according to the types of transitions was carried out. Turns appear to be less accurately detected than transitions between slow/steady regimes and active regimes (A3/A4/A5 to W4 for example, see Table 1 for the categories of changes). This is due to the fact that these turns are not performed the same way by all subjects. Some are too fast to produce a particular pattern in the spectrogram used by our method. This may explain the few errors observed in the segmentation. The segmentation is currently performed using spectrograms whose hop size is 0.1 s, thus limiting the temporal resolution that can be achieved. In case we would like to lower it, it is possible to do so, but at the expense of the computation time.

As for the joint segmentation/classification assessment, the walking phases are well-predicted (>90%) and sufficiently discriminated for the two other phases (>75% for non-sedentary phases and >83% for sedentary phases). Again, imprecision margins are consequences of errors in the segmentation. In addition to segmentation error, inaccuracy can be introduced by the two classifiers. Walking regimes are well detected because this activity is structured and made up of repetitive and precise patterns that therefore manifest themselves with intense spectral signatures. Since a significant proportion of the features used for classification are spectral features, this probably facilitates the classification process. Sedentary and non-sedentary activities are inherently more difficult to differentiate. For example, activities where subjects open fire doors (A1) include movements similar to those observed during walking (stomping and some slow steps) that have spectral signatures closer to those of walking activities. As a result, non-sedentary activities are often mistaken for walking regimes. On the contrary, the intensity of sedentary activities tends to be very low, which may lead to confusion with some low-energy non-sedentary activities. All of these confusions are often encountered in other studies aimed at classifying activities and especially walking activities [52].

#### 4.2. Robustness of the Features

In the final graphical feedback, four features are used to characterize the gait activity. The robustness of the feedback depends mainly on the robustness of those features, especially when confronted with segmentation errors. To investigate this issue, we conducted an additional experiment where we intentionally degrade the segmentation process (e.g., by voluntarily lowering the number of samples for feature computation), in order to assess the robustness of the features. In total, 10 degraded configurations are tested, as described in Table 5. Figure 7 shows the distribution of features in all categories over the 10 configurations in walking sections of one healthy subject (*HSU*) and two pathological subjects: *PSU1* (gluteus medius deficiency) and *PSU2* (post-radiation left brachial plexitis). *PSU1* has shown the highest instability and lack of symmetry in their deambulation, *PSU2* has shown degraded sturdiness. For each subject, we have extracted all walking regimes, and computed the features according to the different configurations. Each box contains the distribution of the different values of this feature on all 10 tested configurations in a given walking regime. The walking regimes for *HSU* are displayed in blue, and the ones of the first, second, and third pathological subject, respectively, in red and green. The blue horizontal line shows the average value of the feature computed on all 10 degraded ranges of every walking regime from all healthy participants (dotted lines correspond to the 75th/25th percentiles). Thus, the distribution for one healthy subject and two pathological subjects as well as the average distribution for all healthy participants are shown in Figure 7.





**Figure 7.** Evaluation of the robustness of selected features. Features with low dispersion and high discrimination between classes. The blue horizontal line shows the average value of the feature for all healthy subjects, the red horizontal line shows the median value of the feature for all healthy subjects, and the dotted lines correspond to the 75th/25th percentiles. Each boxplot corresponds to 10 computations of the feature on a walking regime on 10 degraded ranges. Boxplots are displayed with specific colors depending on their associated subject: blue for a healthy subject, red for *PSU1*, and green for *PSU2*.

One first observation is that all boxes display little spread over all recorded subjects, which suggests that the computation process is robust. It is interesting to note that the differences between the three subjects are clearly visible for all walking regimes in all categories except for steadiness (no recorded participants displayed an affected regularity). Moreover, the patient with the most impact on their stability (*PSU1*) displays boxes associated with the  $RMSR_{ML}$  feature that are even more detached than *PSU2* from the figures of the healthy subjects, which confirms the different visual impacts observed on the gait of each pathological subject. This feature thus presents satisfactory robustness results in terms of dispersion on the degradation ranges as well as in terms of discrimination between subjects. This confirms the relevance of using this feature to evaluate the stability of walking regimes. The calculation of this feature remains indeed constant on all the ranges presented in Table 5, which allows our method to be correctly applied despite eventual segmentation errors that may occur. No patient presented a continuous affection in steadiness and it was thus difficult to estimate the discrimination power from  $P2_{CC}$ : it must be evaluated in further works. Other figures and additional experiments show that all other features listed in Table 3 display the same consistency and robustness, which is an important asset of our proposed approach.

**Table 5.** Degraded configurations for the computation of the features.

Configurations
All the regime is used (normal configuration)
Only the first 3 s of the regime are used
Only the first 3.5 s of the regime are used
Only the first 4 s of the regime are used
Only the first 5 s of the regime are used
Only the first 40% of the regime is used
Only 40% of the regime is used (with start at 20% of the total duration)
Only 40% of the regime is used (with start at 30% of the total duration)
Only 40% of the regime is used (with start at 40% of the total duration)
Only the last 40% of the regime is used

#### 4.3. Relevance of the Graphical Feedback and Possible Usecases

Several important conclusions can be made from the graphical feedbacks obtained on the HSU, PSU1, PSU2, and PSU3 subjects.

*The overall readings of the graphs from HSU in Figure 5a, PSU1 in Figure 5b, and PSU2 in Figure 5c correlate with field observations of subjects' deambulations made by an orthopedic surgeon and a neurologist (Hôpital d'Instruction des Armées Percy):* PSU1 suffers from gluteus medius insufficiency causing Trendelenburg-type lameness because of multiple right hip surgeries. This lameness is related to the low scores obtained for symmetry and stability for this participant. As for PSU2, this patient suffered from post-radiation left brachial plexitis 20 years after radiation treatment for breast cancer. Complete paralysis of the entire left upper limb was observed. The patient's stability and sturdiness are accordingly affected. The displayed impairment of stability may be due to the imbalance related to the dead weight of her left arm, which hangs from the shoulder and weighs at least 10 kg. Besides, PSU2 performed steps with little amplitude which explains degraded scores in sturdiness as displayed in the visual feedback (third inner circle). This confirms that this graphical tool gives a correct overall perception of the subjects' gait actual defects (this must nevertheless be extended to other subjects).

*Longitudinal follow-ups allowed by these visual feedbacks accurately reflect observed degradations or improvements in subjects' walking regimes:* the differences of the same subject before and after their knee ligamentoplasty surgery are presented in Figure 6. In this figure, an obvious degradation is observable between the two recordings as detailed in Section 2.4. This corresponds to the visual observation made by the surgeon: the subject presented a much more degraded gait after surgery than before. After the surgery, there was a significant quadricipital sideration which fully explains the alteration in stability and symmetry (first and fourth inner circles in graphical outputs). This shows how efficient this graphical tool is to track a patient's physical activity longitudinally.

*Visual feedbacks provide time scales of segmented regimes and allow for a new type of ambulatory gait analysis:* regimes' delineations implemented with black lines described in Section 2.5 allow the relative lengths of each segmented regimes to be assessed and compared with each other. This provides a new prism for innovative macro analysis when associated with the evaluations that these visual feedbacks offer: long sedentary regimes might induce better stability scores for walking regimes starting afterward since it can remove fatigue symptoms, for instance.

To the authors' knowledge, few studies based on the use of IMUs in FLEs have endeavored to provide a macro analysis displayed in the form of an easy-to-understand visual legend that fully assesses the entire timeline of FLE signals. Provided summaries either focus on metrics that are too specific, which prevents clear and didactic visual feedback, or on metrics that are too general, which prevent a complete assessment of a subject's physical activity in free environments. In these kinds of studies, the influence of time is often erased by computing features that are often agglomerated over the whole of the measurements as explained above. Rather than knowing the percentage of time spent on each activity, we can, for example, be interested in the impact of transitions between activities on the quality of walking regimes, on the evolution of this quality over several consecutive regimes. In this section, we developed a graphical tool by circumventing these pitfalls of physical activity assessment: time-scales of segmented regimes are provided as well as a continuous evaluation of the physical activity over several walking regimes thanks to pre-defined criteria. Our graphical tool will enable refined follow-ups, displaying an enhanced macro analysis of gait phases. The output graphs allow for easier and more meaningful intra- and interindividual comparisons than those allowed by the global monitoring metrics generally used in the literature. This graphical tool could also allow practitioners to quickly determine the areas of instability in their patients, to identify the influences of fatigue, and perform longitudinal follow-ups allowing new interpretations. Besides, it could help to evaluate rehabilitation procedures or treatment choices for specific

diseases and to assess the impact of treatments on pathologies such as musculoskeletal tumors of the lower limbs or neurological disorders (Parkinson's disease, for instance)

## 5. Conclusions

A pipeline aiming to provide practitioners with a graphical evaluation of their subjects walking in semi-FLEs is presented here. These methods are innovative in the interpretations they offer, adapting a notably concerning segmentation which follows the annotation strategies, precise and ergonomic in its final visual rendering. The rendered visual summary will help practitioners to provide a reliable comprehensive longitudinal tracking of locomotion in free environments. This could for instance allow one to improve post-operative follow-ups and evaluate rehabilitation procedures or treatment choices for specific diseases. The first results are encouraging since they correlate closely with field observations of the walking state of the recorded subjects. Besides, additional criteria such as time scales can enable enhanced interpretations (evaluation of fatigue's impact, development of transitions between activities). However, these methods need to be tested on longer signals in order to define whether they can be applied with the same efficiency to signals collected in FLEs.

**Author Contributions:** Conceptualization, S.J., L.O., and D.R.; data curation, S.J. and N.d.l.; funding acquisition, E.D. and L.G.; investigation, S.J., N.d.l., and L.O.; methodology, S.J., L.O., C.T., and E.D.; software, S.J. and C.T.; supervision, L.O., E.D., L.G., and D.R.; validation, N.d.l. and D.R.; writing—original draft, S.J. and L.O.; writing—review and editing, N.d.l. and D.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki. This experimental protocol was approved by the committee for the protection of individuals (Comité de Protection des Personnes) from the Agence Régionale de Santé (ARS). The ID-RCB number of the committee in which this study is included is: 2021-A00087-34.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data generated during this study are not publicly available due to ethical restriction regarding subjects' personal information but are partially available from the corresponding author upon reasonable request.

**Acknowledgments:** We would like to thank Mona Michaud (Sysnav) for her precious help.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

FLE	Free-Living Environment
Semi FLE	Semi Free-Living Environment
HAR	Human Activity Recognition
IMU	Inertial Measurement Unit

## References

1. Silva de Lima, A.L.; Evers, L.J.; Hahn, T.; Bataille, L.; Hamilton, J.L.; Little, M.A.; Okuma, Y.; Bloem, B.R.; Faber, M.J. Freezing of gait and fall detection in Parkinson's disease using wearable sensors: A systematic review. *J. Neurol.* **2017**, *264*, 1642–1654. <https://doi.org/10.1007/s00415-017-8424-0>.
2. Dot, T.; Quijoux, F.; Oudre, L.; Vienne-Jumeau, A.; Moreau, A.; Vidal, P.P.; Ricard, D. Non-Linear Template-Based Approach for the Study of Locomotion. *Sensors* **2020**, *20*, 1939. <https://doi.org/10.3390/s20071939>.
3. Vienne, A.; Barrois, R.P.; Buffat, S.; Ricard, D.; Vidal, P.P. Inertial sensors to assess gait quality in patients with neurological disorders: A systematic review of technical and analytical challenges. *Front. Psychol.* **2017**, *8*, 817. <https://doi.org/10.3389/fpsyg.2017.00817>.
4. Oudre, L.; Barrois-Müller, R.; Moreau, T.; Truong, C.; Vienne-Jumeau, A.; Ricard, D.; Vayatis, N.; Vidal, P.P. Template-based step detection with inertial measurement units. *Sensors* **2018**, *18*, 4033. <https://doi.org/10.3390/s18114033>.

5. Semwal, V.B.; Gaud, N.; Lalwani, P.; Bijalwan, V.; Alok, A.K. Pattern identification of different human joints for different human walking styles using inertial measurement unit (IMU) sensor. *Artif. Intell. Rev.* **2022**, *55*, 1149–1169.
6. McGrath, T.; Stirling, L. Body-worn IMU-based human hip and knee kinematics estimation during treadmill walking. *Sensors* **2022**, *22*, 2544.
7. Nouredanesh, M.; Godfrey, A.; Howcroft, J.; Lemaire, E.D.; Tung, J. Fall risk assessment in the wild: A critical examination of wearable sensor use in free-living conditions. *Gait Posture* **2021**, *85*, 178–190. <https://doi.org/10.1016/j.gaitpost.2020.04.010>.
8. Halliday, S.J.; Shi, H.; Brittain, E.L.; Hemnes, A.R. Reduced free-living activity levels in pulmonary arterial hypertension patients. *Pulm. Circ.* **2018**, *9*, 2045894018814182. <https://doi.org/10.1177/2045894018814182>.
9. Brodie, M.A.; Coppens, M.J.; Lord, S.R.; Lovell, N.H.; Gschwind, Y.J.; Redmond, S.J.; Del Rosario, M.B.; Wang, K.; Sturnieks, D.L.; Persiani, M.; et al. Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different. *Med. Biol. Eng. Comput.* **2016**, *54*, 663–674. <https://doi.org/10.1007/s11517-015-1357-9>.
10. Jung, S.; Michaud, M.; Oudre, L.; Dorveaux, E.; Gorintin, L.; Vayatis, N.; Ricard, D. The Use of Inertial Measurement Units for the Study of Free Living Environment Activity Assessment: A Literature Review. *Sensors* **2020**, *20*, 5625. <https://doi.org/10.3390/s20195625>.
11. Storm, F.A.; Nair, K.; Clarke, A.J.; Van der Meulen, J.M.; Mazzà, C. Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS ONE* **2018**, *13*, e0196463. <https://doi.org/10.1371/journal.pone.0196463>.
12. Nazarahari, M.; Rouhani, H. Detection of daily postures and walking modalities using a single chest-mounted tri-axial accelerometer. *Med. Eng. Phys.* **2018**, *57*, 75–81. <https://doi.org/10.1016/j.medengphy.2018.04.008>.
13. Cajamarca, G.; Rodríguez, I.; Herskovic, V.; Campos, M.; Riofrío, J.C. StraightenUp+: Monitoring of posture during daily activities for older persons using wearable sensors. *Sensors* **2018**, *18*, 3409. <https://doi.org/10.3390/s18103409>.
14. Ahmadi, M.; O’Neil, M.; Fragala-Pinkham, M.; Lennon, N.; Trost, S. Machine learning algorithms for activity recognition in ambulant children and adolescents with cerebral palsy. *J. NeuroEng. Rehabil.* **2018**, *15*, 105. <https://doi.org/10.1186/s12984-018-0456-x>.
15. Ellis, K.; Kerr, J.; Godbole, S.; Staudenmayer, J.; Lanckriet, G. Hip and wrist accelerometer algorithms for free-living behavior classification. *Med. Sci. Sport. Exerc.* **2016**, *48*, 933–940. <https://doi.org/10.1249/mss.0000000000000840>.
16. Ni, Z.; Wu, T.; Wang, T.; Sun, F.; Li, Y. Deep multi-branch two-stage regression network for accurate energy expenditure estimation with ECG and IMU data. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 3224–3233.
17. Nouredanesh, M.; Tung, J. IMU, sEMG, or their cross-correlation and temporal similarities: Which signal features detect lateral compensatory balance reactions more accurately? *Comput. Methods Programs Biomed.* **2019**, *182*, 105003. <https://doi.org/10.1016/j.cmpb.2019.105003>.
18. Choi, A.; Kim, T.H.; Yuhai, O.; Jeong, S.; Kim, K.; Kim, H.; Mun, J.H. Deep learning-based near-fall detection algorithm for fall risk monitoring system using a single inertial measurement unit. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 2385–2394.
19. Furtado, S.; Godfrey, A.; Del Din, S.; Rochester, L.; Gerrand, C. Free-living monitoring of ambulatory activity after treatments for lower extremity musculoskeletal cancers using an accelerometer-based wearable—A new paradigm to outcome assessment in musculoskeletal oncology? *Disabil. Rehabil.* **2022**, 1–10. <https://doi.org/10.1080/09638288.2022.2083701>.
20. Tawaki, Y.; Nishimura, T.; Murakami, T. Monitoring of gait features during outdoor walking by simple foot mounted IMU system. In Proceedings of the 46th Annual Conference of the IEEE Industrial Electronics Society, IECON, Singapore, 18–21 October 2020; pp. 3413–3418. <https://doi.org/10.1109/IECON43393.2020.9254427>.
21. Oudre, L.; Jakubowicz, J.; Bianchi, P.; Simon, C. Classification of periodic activities using the Wasserstein distance. *IEEE. Trans. Biomed.* **2012**, *59*, 1610–1619. <https://doi.org/10.1109/TBME.2012.2190930>.
22. Truong, C. Détection de Ruptures Multiples—Application aux Signaux Physiologiques. Ph.D. Thesis, Université Paris-Saclay, Paris, France, 2018.
23. Nguyen, M.D.; Mun, K.R.; Jung, D.; Han, J.; Park, M.; Kim, J.; Kim, J. IMU-based spectrogram approach with deep convolutional neural networks for gait classification. In Proceedings of the International Conference on Consumer Electronics (ICCE), Online, 4–6 January 2020; pp. 1–6. <https://doi.org/10.1109/ICCE46568.2020.9042999>.
24. Rehman, R.Z.U.; Klocke, P.; Hryniv, S.; Galna, B.; Rochester, L.; Del Din, S.; Alcock, L. Turning detection during gait: Algorithm validation and influence of sensor location and turning characteristics in the classification of parkinson’s disease. *Sensors* **2020**, *20*, 5377. <https://doi.org/10.3390/s20185377>.
25. Nguyen, H.; Lebel, K.; Bogard, S.; Goubault, E.; Boissy, P.; Duval, C. Using inertial sensors to automatically detect and segment activities of daily living in people with Parkinson’s disease. *IEEE Trans. Neural Syst.* **2017**, *26*, 197–204. <https://doi.org/10.1109/TNSRE.2017.2745418>.
26. Truong, C.; Oudre, L.; Vayatis, N. Greedy kernel change-point detection. *IEEE Trans. Signal Process.* **2019**, *67*, 6204–6214. <https://doi.org/10.1109/TSP.2019.2953670>.
27. Karantonis, D.M.; Narayanan, M.R.; Mathie, M.; Lovell, N.H.; Celler, B.G. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Trans. Inf. Technol.* **2006**, *10*, 156–167. <https://doi.org/10.1109/titb.2005.856864>.
28. Killick, R.; Fearnhead, P.; Eckley, I.A. Optimal detection of changepoints with a linear computational cost. *J. Am. Stat. Assoc.* **2012**, *107*, 1590–1598. <https://doi.org/10.1080/01621459.2012.737745>.

29. Truong, C.; Oudre, L.; Vayatis, N. Penalty learning for changepoint detection. In Proceedings of the 25th European Signal Processing Conference (EUSIPCO), Kos Island, Greece, 28 August–2 September 2017; pp. 1569–1573. <https://doi.org/10.23919/EUSIPCO.2017.8081473>.
30. Jung, S.; Oudre, L.; Truong, C.; Dorveaux, E.; Gorintin, L.; Vayatis, N.; Ricard, D. Adaptive change-point detection for studying human locomotion. In Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Guadalajara, Mexico, 26 July 2021; pp. 2020–2024. <https://doi.org/10.1109/embc46164.2021.9629775>.
31. Prasanth, H.; Caban, M.; Keller, U.; Courtine, G.; Ijspeert, A.; Vallery, E.; Von Zitzewitz, J. Wearable sensor-based real-time gait detection: A systematic review. *Sensors* **2021**, *21*, 2727.
32. Rast, F.M.; Labruyère, R. Systematic review on the application of wearable inertial sensors to quantify everyday life motor activity in people with mobility impairments. *J. Neuroeng.* **2020**, *17*, 1–19. <https://doi.org/10.1186/s12984-020-00779-y>.
33. Fiorini, L.; Bonaccorsi, M.; Betti, S.; Esposito, D.; Cavallo, F. Combining wearable physiological and inertial sensors with indoor user localization network to enhance activity recognition. *J. Ambient Intell. Smart Environ.* **2018**, *10*, 345–357. <https://doi.org/10.3233/AIS-180493>.
34. Kerr, J.; Patterson, R.E.; Ellis, K.; Godbole, S.; Johnson, E.; Lanckriet, G.; Staudenmayer, J. Objective assessment of physical activity: Classifiers for public health. *Med. Sci. Sport. Exerc.* **2016**, *48*, 951–957. <https://doi.org/10.1249/MSS.0000000000000841>.
35. Marcotte, R.T.; Petrucci Jr, G.J.; Cox, M.F.; Freedson, P.S.; Staudenmayer, J.W.; Sirard, J.R. Estimating Sedentary Time from a Hip-and Wrist-Worn Accelerometer. *Med. Sci. Sport Exerc.* **2020**, *52*, 225–232. <https://doi.org/10.1249/MSS.0000000000002099>.
36. Fullerton, E.; Heller, B.; Munoz-Organero, M. Recognizing Human Activity in Free-Living Using Multiple Body-Worn Accelerometers. *IEEE Sens. J.* **2017**, *17*, 5290–5297. <https://doi.org/10.1109/JSEN.2017.2722105>.
37. Hsu, C.Y.; Tsai, Y.S.; Yau, C.S.; Shie, H.H.; Wu, C.M. Differences in gait and trunk movement between patients after ankle fracture and healthy subjects. *Biomed. Eng. Online* **2019**, *18*, 1–13. <https://doi.org/10.1186/s12938-019-0644-3>.
38. Sekine, M.; Tamura, T.; Yoshida, M.; Suda, Y.; Kimura, Y.; Miyoshi, H.; Kijima, Y.; Higashi, Y.; Fujimoto, T. A gait abnormality measure based on root mean square of trunk acceleration. *J. Neuroeng.* **2013**, *10*, 1–7. <https://doi.org/10.1186/1743-0003-10-118>.
39. Bahari, H.; Forero, J.; Hall, J.C.; Hebert, J.S.; Vette, A.H.; Rouhani, H. Use of the extended feasible stability region for assessing stability of perturbed walking. *Science* **2021**, *11*, 1026. <https://doi.org/10.1038/s41598-020-79955-y>.
40. Bruijn, S.M.; Meijer, O.; Beek, P.; van Dieen, J.H. Assessing the stability of human locomotion: A review of current measures. *J. R. Soc. Interface* **2013**, *10*, 20120999. <https://doi.org/10.1098/rsif.2012.0999>.
41. Ben Mansour, K.; Gorce, P.; Rezzoug, N. The Multifeature Gait Score: An accurate way to assess gait quality. *PLoS ONE* **2017**, *12*, e0185741. <https://doi.org/10.1371/journal.pone.0185741>.
42. Labaune, O.; Deroche, T.; Teulier, C.; Berret, B. Vigor of reaching, walking, and gazing movements: On the consistency of interindividual differences. *J. Neurophysiol.* **2020**, *123*, 234–242. <https://doi.org/10.1152/jn.00344.2019>.
43. Kobayashi, H.; Kakihana, W.; Kimura, T. Combined effects of age and gender on gait symmetry and regularity assessed by autocorrelation of trunk acceleration. *J. Neuroeng.* **2014**, *11*, 1–6. <https://doi.org/10.1186/1743-0003-11-109>.
44. Moe-Nilssen, R.; Helbostad, J.L. Estimation of gait cycle characteristics by trunk accelerometry. *J. Biomech.* **2004**, *37*, 121–126. [https://doi.org/10.1016/S0021-9290\(03\)00233-1](https://doi.org/10.1016/S0021-9290(03)00233-1).
45. Tura, A.; Raggi, M.; Rocchi, L.; Cutti, A.G.; Chiari, L. Gait symmetry and regularity in transfemoral amputees assessed by trunk accelerations. *J. Neuroeng.* **2010**, *7*, 1–10. <https://doi.org/10.1186/1743-0003-7-4>.
46. Cabral, S. Gait Symmetry Measures and Their Relevance to Gait Retraining. In *Handbook of Human Motion*; Springer International Publishing: Cham, Switzerland, 2018; pp. 429–447. [https://doi.org/10.1007/978-3-319-14418-4\\_201](https://doi.org/10.1007/978-3-319-14418-4_201).
47. Brard, R.; Bellanger, L.; Chevreuil, L.; Doistau, F.; Drouin, P.; Stamm, A. A novel walking activity recognition model for rotation time series collected by a wearable sensor in a free-living environment. *Sensors* **2022**, *22*, 3555.
48. Garcia-Gonzalez, D.; Rivero, D.; Fernandez-Blanco, E.; Luaces, M.R. A public domain dataset for real-life human activity recognition using smartphone sensors. *Sensors* **2020**, *20*, 2200.
49. Mo, L.; Zhu, Y.; Zeng, L. A Multi-Label Based Physical Activity Recognition via Cascade Classifier. *Sensors* **2023**, *23*, 2593.
50. Cescon, M.; Choudhary, D.; Pinsker, J.E.; Dadlani, V.; Church, M.M.; Kudva, Y.C.; Doyle III, F.J.; Dassau, E. Activity detection and classification from wristband accelerometer data collected on people with type 1 diabetes in free-living conditions. *Comput. Biol. Med.* **2021**, *135*, 104633.
51. Konsolakis, K. Physical Activity Recognition Using Wearable Accelerometers in Controlled and Free-Living Environments. Master's Thesis, Delft University of Technology, Delft, The Netherlands, 2018.
52. Andreu-Perez, J.; Garcia-Gancedo, L.; McKinnell, J.; Van der Drift, A.; Powell, A.; Hamy, V.; Keller, T.; Yang, G.Z. Developing fine-grained actigraphies for rheumatoid arthritis patients from a single accelerometer using machine learning. *Sensors* **2017**, *17*, 2113. <https://doi.org/10.3390/s17092113>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.