UNIVERSITÀ
DEGLI STUDI
FIRENZE

PhD Program in Smart Computing
Dipartimento di Ingegneria dell'Informazione (DINFO)

Assessment of Frailty using a Wrist-worn Device

Domenico Minici

Dissertation presented in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Smart Computing

PhD Program in Smart Computing
University of Florence, University of Pisa, University of Siena

Assessment of Frailty using a Wrist-worn Device

Domenico Minici

Advisors:

Prof. Marco Avvenuti
Dr. Guglielmo Cola

Head of the PhD Program:

Prof. Stefano Berretti

Evaluation Committee:
Prof. Jeff D. Williamson, *Wake Forest University*
Dr. Daniele Ravì, *University of Hertfordshire*

XXXV ciclo — October 2022

To my Lilli

Abstract

The COVID-19 pandemic has considerably shifted the focus of scientific research, speeding up the process of digitizing medical monitoring. Wearable technology is already widely used in medical research, as it has the potential to monitor the user's physical activity in daily life. Therefore, they are particularly appealing for evaluating older subjects in their environment to capture early signs of frailty and mobility-related problems. Early detection of abnormal physical performance and gait may help identify physically frail subjects at an increased risk of losing their independence but are still amenable to preventive interventions, such as structured exercise.

This dissertation explores the use of body-worn accelerometers for automated assessment of frailty during walking activity. We proposed an automated process based on machine learning techniques, able to classify patients according to their frailty status using a set of gait-related parameters extracted from wearable sensors. Here, we highlight the importance of the position chosen for placing the sensors by comparing the performances of a wrist- against a lower back-worn sensor.

Secondly, the dissertation analyzes the use of Continuous Wavelet Transform in combination with sensor-derived gait parameters for frailty status assessment. Continuous Wavelet Transform was applied to obtain time-frequency domain representations of gait signals. Here, the most statistically significant band-based information for frailty status assessment was identified by means of ANOVA and statistical t-test. Moreover, a Deep Convolutional Neural Network was trained and tested for identifying wavelet-based patterns to classify subjects as robust or non-robust, a category that includes both Fried's frail and pre-frail phenotypes.

Finally, the dissertation aims to explore in-home collected wearable-derived signals for frailty status assessment. Signal-derived traces were segmented. A subset of these segments was used to calculate the Subject Activity Level, an index to quantify how users were active throughout the day. The SAL index was then combined with gait-derived features to design a novel frailty status assessment algorithm.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

It has been documented that, besides its direct, dramatic effects on subjects suffering from the disease, the COVID-19 outbreak has been responsible for severely decreased care opportunities in other clinical conditions because the need for social distancing restricted direct contact between patients and their healthcare providers (Hantke and Gould, 2020). Consequently, healthcare systems are nowadays increasingly trying to develop new paradigms of care, which would value enhanced remote clinical assessment and monitoring, as well as the delivery of therapies at a distance. Different systems for remote assessment and care have already gained considerable attention, particularly for patients in isolated communities and remote regions (Albahri et al., 2018; Sanders et al., 2013). Older persons with chronic comorbidities, frequently unable to reach the assessment clinic, are a further important target of such efforts for the development of innovative care technologies, because they were proven to be particularly affected by the restriction in the application of the conventional model of care during the COVID-19 pandemic, with strong, negative impacts (Gilpin, 2018; Hantke and Gould, 2020).

According to a World Health Organization (WHO) estimate, individuals over the age of 60 will nearly double over the 2015-2050 period, increasing from 12% to 22% of the world population. Moreover, 80% of the older adults have at least one chronic disease, and 77% have at least two [1]. Among the conditions that undermine the quality of life of older persons, frailty represents one of the most severe global public health challenges (Dent et al., 2019).

In recent years, clinical gerontology has devoted a growing interest to frailty. Frailty is a condition in which individuals are highly vulnerable to both endogenous and external stressors, exposing them to an excess risk of poor clinical outcomes, such as falls and fractures, disability, hospitalization, and ultimately death (Collard et al., 2012). Not surprisingly, frail older subjects absorb a significant share

[1] *Mental health of older adults*, `https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults`

of healthcare resources (Fairhall et al., 2015).

One aspect often overlooked is represented by the economic impact of frailty on global healthcare costs. The European Commission has produced a report to foresee long-term care costs, where long-term care is usually defined as a set of services required by a person with a reduced degree of functional capacity, physical or cognitive. In the European Union, an increase from the actual 1.6% to 2.7% of Gross Domestic Product in 2070 is expected [2]. The increasing costs will represent a massive challenge for the sustainability of long-term care, and frailty is seen as a major factor in the loss of functional independence. Other studies have explored the impact of frailty on health care costs (Bock et al., 2016; Sirven and Rapp, 2017; Garcìa-Nogueras et al., 2017; Simpson et al., 2018; Salinas-Rodrìguez et al., 2019). As diverse as these studies may be, in the method and sample used, they arrive at the same conclusion: there is a strong correlation between the increase in frail patients and the increase in public health-related costs.

Within the general definition given above, diverse conceptual models and, consequently, different diagnostic tools have been proposed for frailty. One of the most accredited among them is the phenotype model, which was developed by Fried et al. (Fried et al., 2001) and considers a decline in physical performance as the cornerstone of frailty. Although this tool is simple and of rapid application, it requires some specialized clinical setup and trained personnel. Moreover, it might be hypothesized that exploring the physical performance of older subjects in their own environment is more appropriate to capture frailty status than the somewhat artificial setting needed to apply this and other similar tests. For these reasons, many researchers have focused on the use of wearable, sensor-based technology to gather parameters on motor condition (Dasenbrock et al., 2016; Schwenk et al., 2015; Mohler et al., 2014; Abbate et al., 2012), thus obtaining an objective, ecological assessment of frailty status in older subjects (Schwenk et al., 2014; Thiede et al., 2016). Furthermore, patient monitoring through wearable technology can enable prolonged studies, thus expanding the bulk of data available for evaluation, while reducing healthcare costs and the discomfort for the patient when the assessment is done within a specific clinical setting.

In the following subsections, we will briefly discuss this Ph.D. dissertation's aims and objectives to address the aspects described above.

Towards automated assessment of frailty status

This chapter aims to devise an automated process based on machine learning, able to classify subjects according to their frailty status using a set of gait-related parameters

[2] *The 2018 Ageing Report*, `https://ec.europa.eu/info/sites/default/files/economy-finance/ip079_en.pdf`

extracted from wearable accelerometers. Here, we report on the basic approach and the experimental protocol, and we highlight the importance of the position chosen for placing the sensors, comparing the performance of a wrist- and a lower back-worn sensor. We, therefore, propose a method for distinguishing robust (R) subjects from pre-frail and frail subjects, joined under the common denomination of non-robust (NR). This binary classification still allows geriatricians to identify the individuals who may need further clinical evaluation, namely pre-frail and frail. Indeed, early identification of the latter categories enables prompt clinical intervention, which in turn may lead to better health outcomes for the subject.

A three-stage approach is used for frailty status assessment, exploiting either the wrist or the lower-back sensor. In the first stage, gait cycle detection is executed to identify segments of four gait cycles. In the next stage, gait signals are analyzed to extract 25 features characterizing the subject's gait. Finally, the last stage uses machine learning to classify participants as R or NR.

We compared two solutions based on a single device as we aim to provide an unobtrusive solution for frailty assessment.

The experiments to evaluate this approach involved 34 volunteers, who were asked to walk at a preferred speed under the supervision of experienced medical staff. Clinicians were also responsible for the classification of participants according to Fried's frailty phenotype. The results of the clinical assessment were used to label accelerometer signals and built a ground-truth dataset that, in turn, was used to train, validate and test machine learning-based classification of frailty.

The main contributions of this chapter are to (i) gain a better understanding of the possibility of using practical wearable devices in combination with machine learning-based classification to discriminate R from NR older subjects, (ii) show that walk-related features in the time-frequency domain, computed by means of Wavelet analysis, can improve performance in frailty status assessment, (iii) compare the performance of sensors applied in two different body parts, the wrist and the lower back.

Wavelet-based analysis of gait for automated frailty assessment

In this chapter, we aim to improve the assessment of frailty based on gait-related parameters derived from a wrist-worn device. Specifically, we focus on potential predictors offered by the Continuous Wavelet Transform (CWT), which allows us to observe the evolution of frequency-related parameters over time. We believe that a deeper investigation of the CWT frequency content, aimed at identifying the most relevant frequency bands, could be beneficial to the field of automated frailty assessment based on wearable sensors.

A total of 34 volunteers aged 70+ were initially screened by geriatricians for the presence of frailty according to Fried's criteria. After screening, participants were

asked to perform a 60 m walk test at preferred pace, while wearing an accelerometer on the wrist. A gait detection technique was applied to the sensor-derived signal, in order to identify segments made of four gait cycles. Continuous Wavelet Transform was applied to obtain time-frequency domain representations, which were subsequently used in a band-based feature extraction phase. Here, the most significant band-based features for frailty status assessment were identified by means of ANOVA and statistical t-test. Finally, a Random Forest for each frequency band was trained and tested for classifying subjects as robust or non-robust (i.e., pre-frail or frail). This information may help achieve continuous assessment of frailty in older adults with a wrist-worn device.

A deep learning approach for frailty status assessment

This chapter discusses the development of a deep learning-based system to support clinicians in the preventive screening of frailty status assessment. Starting from previous studies and results, we applied a deep learning model to gait signals to automatically classify subjects according to their frailty status. The chapter explores the use of a deep neural network combined with Wavelet Analysis applied to gait accelerometric signals acquired at the wrist to classify robust and non-robust subjects.

The initial dataset was the same used for the previous studies and already presented above. In particular, signals recorded during gait were processed, split into fixed-length windows, and given as input to Continuous Wavelet Transform. Outputs from Wavelet analysis were used to train and test a Convolutional Neural Network to distinguish subjects according to their frailty status.

Automated, ecologic assessment of frailty using a wrist-worn device

This chapter explores the possibility of assessing the frailty status of a subject based on a signal collected through wearable technology in an in-home environment. For this purpose, we aim to demonstrate how an objectively measured activity level might characterize subjects according to their frailty status. In addition, we intend to show how gait biomarkers collected through a wrist-worn device may represent an essential step forward for continuous frailty monitoring.

To this end, subjects aged 70+ years were enrolled and classified into robust, pre-frail, and frail according to Fried's phenotype (Fried et al., 2001). A device equipped with inertial motion sensors was fixed to the subjects' wrists for 24 hours, during which the participants conducted their everyday lives. The accelerometer signal traces collected were split into segments of 10-second duration, which were then analyzed using a previously developed algorithm (Cola et al., 2014) designed to recognize gait segments. All non-gait segments were further analyzed and classified

as other motor activities when they exceeded a pre-determined activity threshold or rest segments. A Subject Activity Level (SAL) was then calculated using the number of gait and other motor activities. This parameter then provided the input to the frailty status assessment algorithm. The algorithm compared the subject's SAL with pre-determined thresholds to classify the patient as robust (R) or non-robust (NR), which included both frail and pre-frail subjects. The gait traces of subjects that cannot be immediately classified by the SAL index alone were further analyzed through a Machine Learning model, which assigns the R or NR class based on the features computed on the gait segments.

Chapter 2

Background

This Ph.D. dissertation explores the use of wearable devices equipped with inertial sensors for frailty identification. In the following sections, we provide a brief background of the topics underlying the research presented in the dissertation's chapters.

2.1 Frailty

Frailty is common in later life across different countries Manfredi et al. (2019); the reported prevalence increases progressively with aging and reaches approximately 30% in subjects aged 85+ years Collard et al. (2012). As pointed out in several studies van Kan et al. (2008); Collard et al. (2012), there is a lack of consensus on the definition of frailty and its clinical identification. A turning point in the conceptualization of frailty was represented by a study by Fried et al., who in 1991 hypothesized the existence of a stage of preclinical disability in which there is a greater risk of functional decline Fried et al. (1991). A few years later, Guralnik et al. Guralnik et al. (1995) demonstrated that measures of lower extremity functioning, obtained with a simple battery of three physical performance tests, might be a hallmark of frailty: in non-disabled subjects, poorer scores on this battery predicted the subsequent development of disability in the two domains of mobility and basic activities of daily living (BADL), thus providing evidence that a state of preclinical disability can be clinically identified Guralnik et al. (1994); Pritchard et al. (2017); Abizanda et al. (2012). Further advancement was offered by Fried et al. in a 2001 landmark study (Fried et al., 2001), where researchers introduced the concept of frailty phenotype, accompanied by the proposal of a simple tool to recognize frailty through the following dimensions:

1. an unintentional weight loss of 4.5 kg or more in the last year;
2. low energy, identified through the CES-D (Center of Epidemiologic Studies Depression Scale) (Orme et al., 1986);

3. low physical activity, defined thanks to the Physical Activity Questionnaire for the Elderly (PASE) (Schuit et al., 1997);

4. slowness, defined by the speed measured over a distance of 4.5 m and normalized for height and gender;

5. weakness, meaning reduced muscle strength in the dominant hand.

A subject is considered: *frail* if positive for three or more dimensions; *pre-frail* if positive for one or two dimensions; *robust* if negative for all dimensions.

A substantial body of literature has shown that older persons identified as frail by these criteria have an increased risk of accelerated functional decline, overt disability, and other poor clinical outcomes, including death. Thus, Fried's frailty phenotype has become the most commonly accepted model to identify frail older subjects Guralnik et al. (1995); Bergman et al. (2003); Visser et al. (2003); Collard et al. (2012). Given the importance of physical functioning and, more specifically, lower extremity mobility in the detection of frailty, it is reasonable to hypothesize that an extended evaluation of mobility may improve our ability to identify this condition more precisely and at an earlier stage.

2.2 Activities of Daily Living

Many longitudinal studies have reported a significant decline in physical performance in frailty, leading to the need for care and support in the Activities of Daily Living (ADL) (Gobbens et al., 2010; Xue, 2011). ADL is a term used to collectively describe fundamental skills required to independently care for oneself, such as eating, bathing, and mobility. The term was first coined by Sidney Katz in 1950 (Edemekong et al., 2021; Katz, 1983; Bieńkiewicz et al., 2014). The relationship between frailty and ADL has been widely discussed in literature (Mañas et al., 2017; García-Esquinas et al., 2017; Dipietro et al., 2019; Kehler et al., 2018; Kehler and Theou, 2019). Dipietro et al. (Dipietro et al., 2019) suggest that regular physical activity effectively helps older adults delay the loss of mobility while reducing the risk of fall-related injuries. Authors in (Kehler and Theou, 2019) conclude that spending less time in sedentary pursuits confers a protective association with frailty. However, the authors highlight the lack of epidemiological studies confirming this. Of the same idea are the authors in (Ziller et al., 2020), whose results emphasize the importance of physical activities and their assessment methods.

2.3 Gait analysis

Among the activities that can be investigated using wearable sensors, gait plays a very important role in identifying age-related conditions. Generally, human walking is a periodic movement of the body segments and includes repetitive motions. To understand this periodic walking course better and easier, the gait phase must be used to describe an entire walking period (Tao et al., 2012).

Gait analysis represents a critical tool for identifying clinical conditions since it can capture declines in physical performances. Gait analysis has been the subject of several studies in the context of personalized healthcare, and gait-related parameters derived from wearable sensors have already been associated with conditions such as frailty (Kosse et al., 2016; Pradeep Kumar et al., 2020).

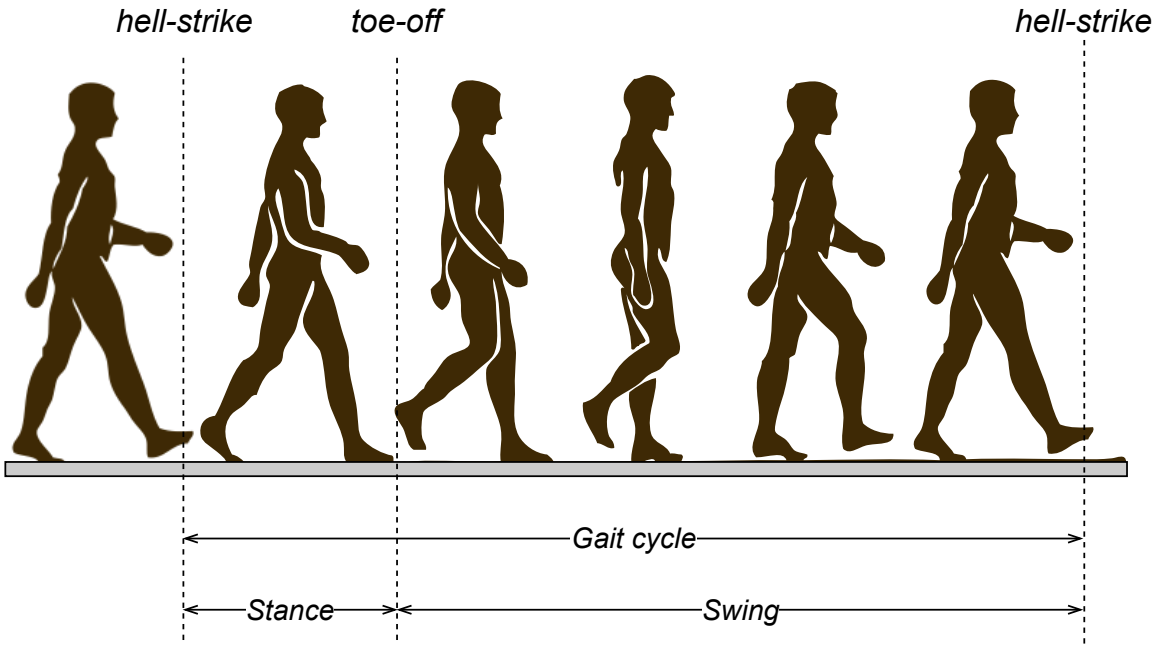


Figure 2.1: Representation of gait events involved in the human walk.

Figure 2.1 shows an example of gait. Two main events can be identified during the human walk: *toe-off* and *hell-strike*. Toe-off is the moment in which the foot loses contact with the ground to be moved forward. On the other hand, heel-strike is the moment in which the foot touches the ground again. A *gait cycle* is the sequence of events that occur during the walking process between two consecutive heel strikes of the same foot.

2.4 Sensor-based frailty identification

In clinical settings, gait is often analyzed using specialized equipment, such as a treadmill with integrated pressure sensors. At the same time, ADLs are usually evaluated using self-reporting or interview-based assessment tools (Pashmdarfard and Azad, 2020). In other words, such measures typically require some specialized clinical setup and trained personnel. Inertial sensors have been proposed as powerful tools to assess functional capacity Millor et al. (2013); Zhang et al. (2015); Avvenuti et al. (2018). These devices allow the collection of a considerable amount of data on subjects' mobility in their home environment over prolonged times; data can then be automatically transmitted at a distance and analyzed with appropriate algorithms, potentially reducing operator-dependent variability in the assessment of frailty.

In particular, accelerometers and gyroscopes have been used to measure mobility parameters or assess the risk of falls in the elderly Zivanovic et al. (2018); Díaz et al. (2020); Cola et al. (2016); Pradeep Kumar et al. (2020). Several researchers studied the parameters recorded by body-worn inertial sensors to discriminate patients according to their frailty status. The authors in Galán-Mercant and Cuesta-Vargas (2013) described the differences between frail and non-frail older subjects using parameters extracted from inertial sensors: in particular, they found that frail elderly persons obtained lower maximum and minimum accelerations than fit individuals. The authors in Schwenk et al. (2015) found that unique parameters derived from an objective assessment of gait, balance, and physical activities are sensitive for identifying pre-frailty and classifying a subject's frailty status. In Greene et al. (2014), Greene et al. concluded that it is possible to distinguish frail from non-frail subjects by using mobility tests in combination with wearable sensors. In (Martínez-Ramírez et al., 2015), Martínez-Ramírez et al. extracted a collection of sensor-based parameters to supply more accurate information about older adults' frailty syndrome. In Ritt et al. (2017), Ritt et al. presented strong evidence about the correlation between gait alteration and frailty using a shoe-mounted inertial- sensor-based mobile gait analysis system. Zhong et al. (Zhong et al., 2018) analyzed smart bracelets' signals to establish reference gait parameters. More recently, authors in (Pradeep Kumar et al., 2020) collected data using a tri-axial accelerometer fixed at the sternum to suggest everyday gait characteristics, along with quantitative measures of physical activity, as an opportunity to screen frailty. Researchers in (García-Villamil et al., 2021) developed a sensor-based system to measure gait parameters in older adults with falls and studied how these parameters correlate to different frailty levels. A relevant project is represented by Frailsafe (Zacharaki et al., 2020), which proposes a framework that highlights the potential for frailty prediction strategies based on information and communication technology (ICT). Both quantitative and qualitative

measures of frailty are used to predict long-term outcomes via advanced data mining approaches applied to multi-parametric data. Frailsafe uses wearable sensors for monitoring several physical activities and physiological parameters.

The precise procedures and methods to apply wearable technology in frailty assessment are still under investigation, and uncertainties exist on the best position of the sensors and the parameters to be extracted, and the algorithm to process them. We indeed argue that the position chosen for wearing the sensor can be of fundamental importance, not only for usability reasons but also for recognizing frailty. For this reason, this dissertation compares two different sensor positions (lower back and wrist). In the literature, other researchers analyzed the importance of data collected from wrist-worn accelerometers in older adults' health, functional, and social assessment. Huisingh-Scheetz et al. used the signal extracted from wrist-worn sensors to determine how frailty and other characteristics relate to activity among older adults Huisingh-Scheetz et al. (2016, 2018). In our case, besides using a machine-learning approach, we focus on the relationship between gait-related movements and the subject's frailty status. Notably, while in many previous works inertial signals are mainly treated as static processes, in the proposed methods, we devote particular attention to the time-frequency domain using Wavelet analysis.

2.5 Continuous Wavelet Transform

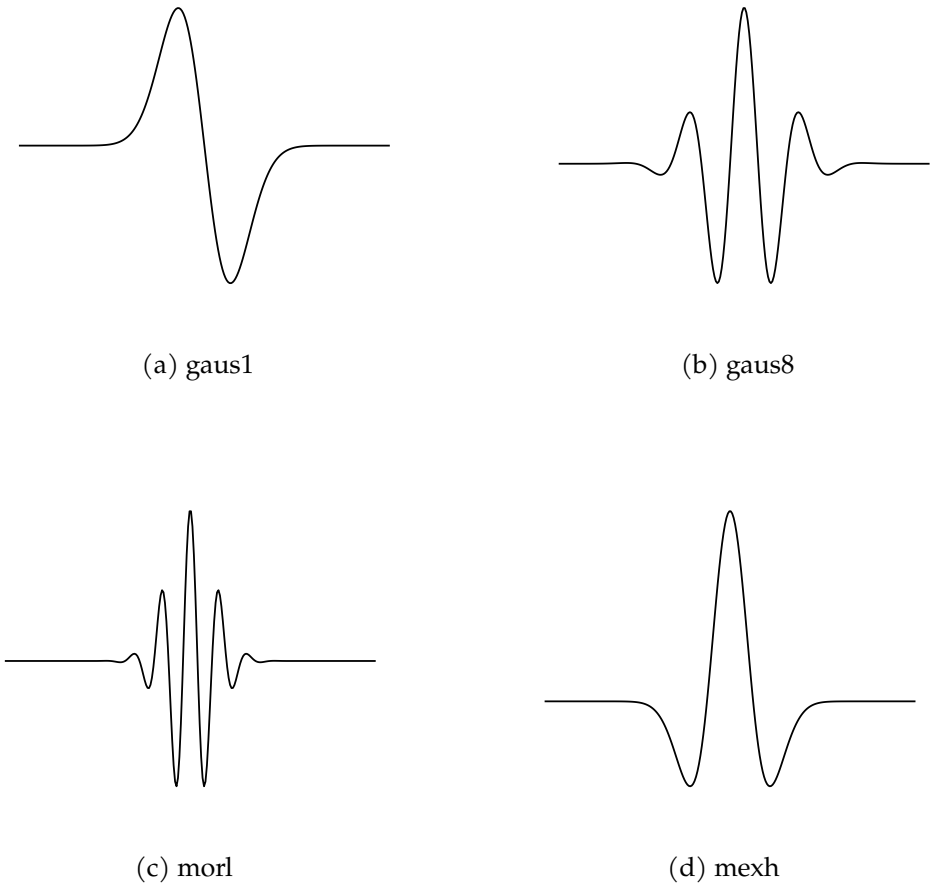
In recent years, other researchers investigated the use of *Continuous Wavelet Transform* (CWT) in combination with sensor-derived data extracted during different activities (Martínez-Ramírez et al., 2011; Khandelwal and Wickström, 2016; Rezvanian and Lockhart, 2016).

The Wavelet Transform method is a signal processing technique for analyzing a time series containing non-stationary power at different frequencies (Daubechies, 1990). In particular, it is possible to analyze local power variations for each component by decomposing a time series into various frequency components. Wavelet theory uses a set of Mother Wavelets (MW), which are scaled and translated to obtain a time-frequency representation of the time series.

In the next chapters, we will use the CWT W_n of a discrete sequence r_n with equal time spacing δt, which is defined as the convolution of r_n with a scaled and translated version of the MW ψ_0 (Torrence and Compo, 1998):

$$W_n(s) = \sum_{n_i=0}^{N-1} r_n \cdot \overline{\psi} \left[\frac{(n_i - n)\delta t}{s} \right],$$

where $\overline{\psi}$ indicates the complex conjugate of the wavelet function, s is the scale factor, $(n_i - n)$ denotes the translation along the gait segment, and δt is the inverse of the sampling rate.



(a) gaus1 (b) gaus8

(c) morl (d) mexh

Figure 2.2: Wavelet families used in Continuous Wavelet Analysis.

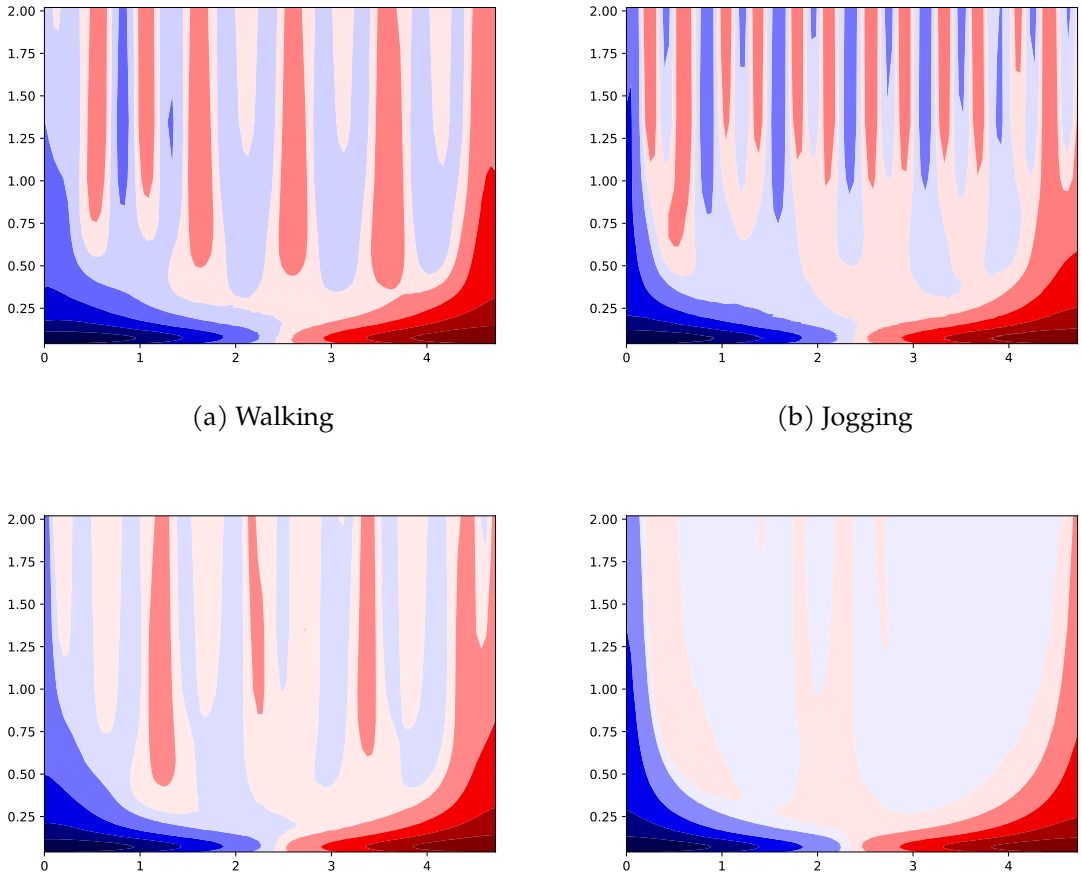
The Mother Wavelet function(ψ_0) plays a key role, it must have zero mean and be localized both in frequency and time domains. Scaling the mother wavelet means stretching or compressing the it through the scaling factor. This mechanism gives information about the frequency domain. On the other hand, translation corresponds to a shift in time and describes the Mother Wavelet's position in the time domain.

Figure 2.2 shows the MW families used in all our experiments: Gaussian (*gaus*), Morlet (*morl*), and Mexican Hat (*mexh*). Let *gausN* indicate a gaussian wavelet with N vanishing moments: N is related to the approximation order and smoothness of the wavelet, so that a wavelet with N vanishing moments can approximate polynomials of degree $N-1$. In this dissertation, $N \in [1, 8]$.

Continuous Wavelet Transform is especially suited for the analysis of the human body acceleration magnitude signal due to its dynamic nature Torrence and Compo (1998).

Some examples of outputs produced by the application of CWT to analyze different human activities with different MW are shown in Figures 2.3, 2.4, and 2.5. These

outputs are called scalograms, and show variations of power over time (**x-axis**) at different frequency bands (**y-axis**). Here, colors represent the intensity of power - from blue (low power) to red (high power).



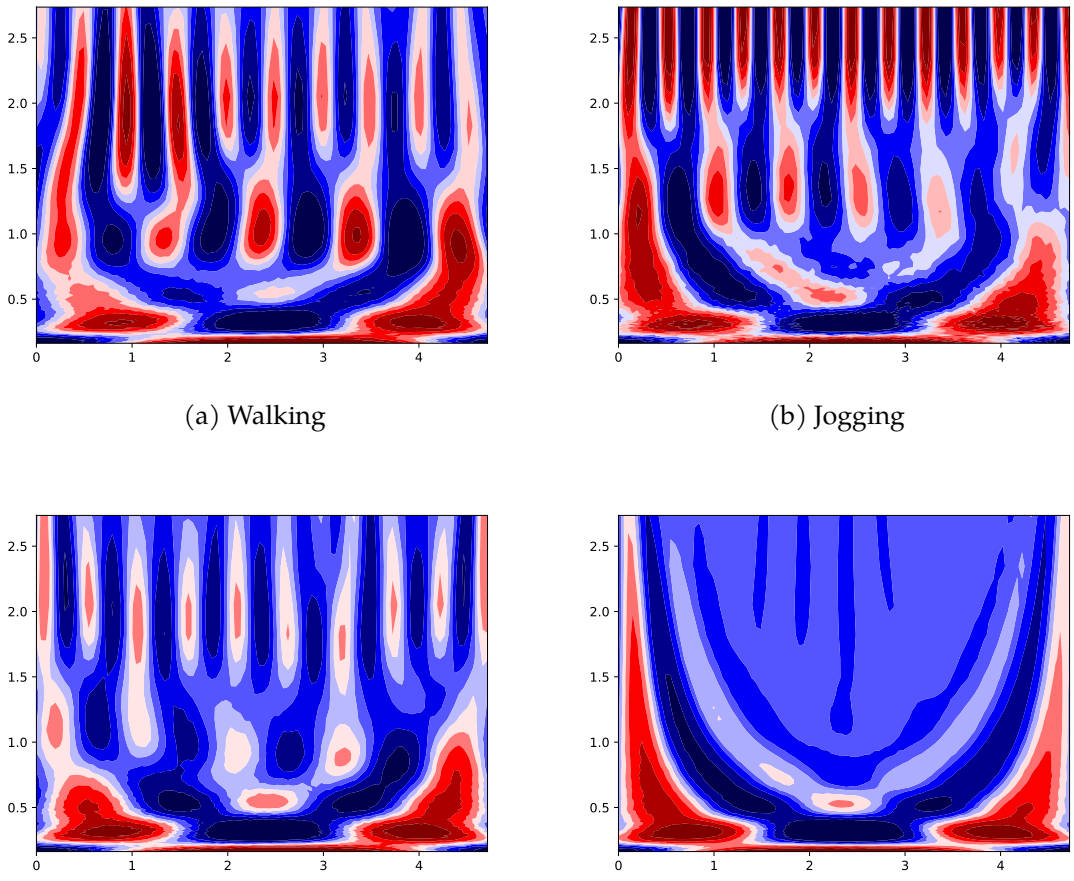
(a) Walking

(b) Jogging

(c) Stairs

(d) Sitting

Figure 2.3: Continuous Wavelet Transform scalograms obtained from the *gaus1* mother wavelet applied to human activity signals.



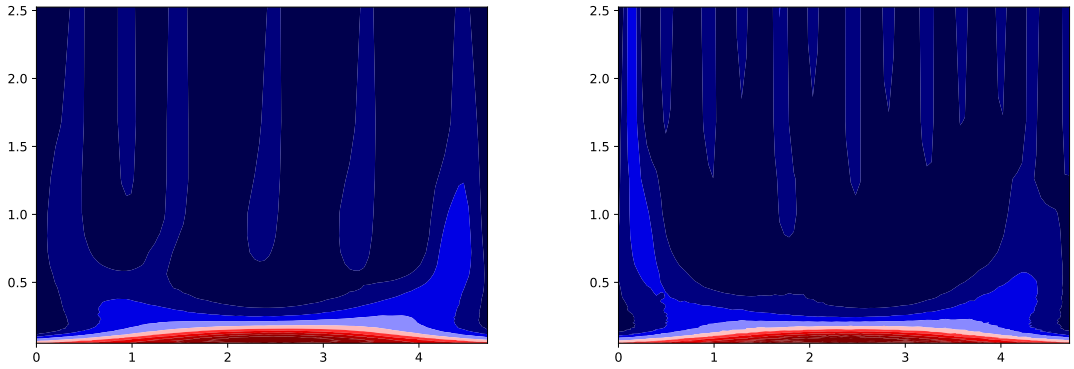
(a) Walking

(b) Jogging

(c) Stairs

(d) Sitting

Figure 2.4: Continuous Wavelet Transform scalograms obtained from the *morl* mother wavelet applied to human activity signals.



(a) Walking

(b) Jogging

(c) Stairs

(d) Sitting

Figure 2.5: Continuous Wavelet Transform scalograms obtained from the *mexh* mother wavelet applied to human activity signals.

Chapter 3

Towards automated assessment of frailty status

The chapter is organized as follows. First, Section 3.1 describes the method we proposed for automatic frailty assessment based on wearable accelerometers. In Section 3.2, the setup of the experiment, the design choices and their validation are presented. Finally, in Section 3.3 the results achieved are presented and discussed.

3.1 The proposed method

In this section, we give a system description of the method we propose for automatic frailty assessment. The flowchart of the method is shown in Figure 3.1. At a glance, acceleration data is collected through a sensor embedded in a wearable device. After preprocessing, acceleration samples are sent to a gait segment detection module, which works by isolating sequences of consecutive gait cycles from the incoming signal trace. Once the gait segments have been detected, a subset containing the most regular walks is selected and used for the feature extraction phase. Finally, extracted features are used to feed a machine learning classifier which assesses whether the bearer is a robust (R) or non-robust (NR) subject.

Data acquisition and preprocessing

Acceleration is collected at a sampling rate of 102.4 Hz. All the acceleration components (**x**, **y**, **z** axes and acceleration magnitude **m**) are converted into g units. As body movements typically have frequency components below 20 Hz (Antonsson and Mann, 1985), acceleration components are passed through a second order Butterworth low-pass filter with a cut-off frequency of 20 Hz. Filtered acceleration is then provided to the gait segment detection module.

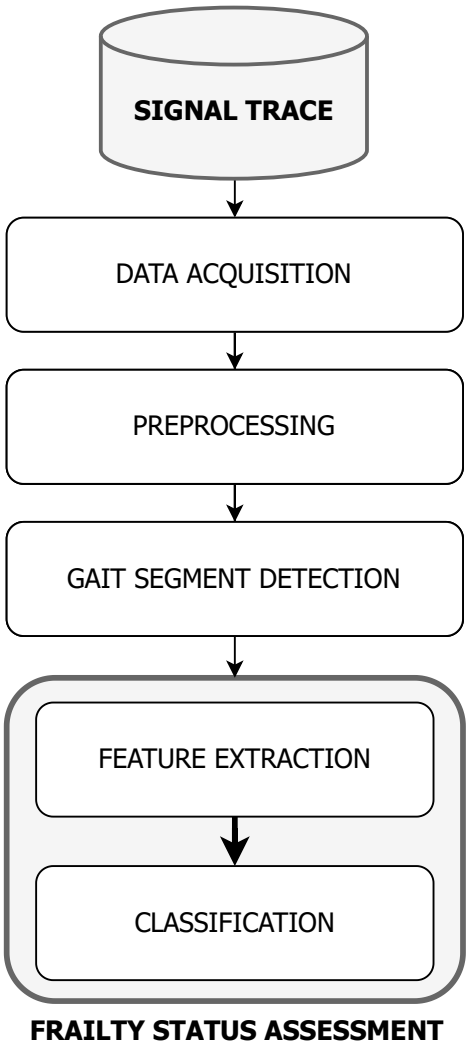


Figure 3.1: Flowchart of the proposed method for frailty status assessment.

The frequency response of the filter is reported in Figure 3.2. Here, the vertical green line corresponds to the cutoff frequency of the filter computed as follows:

$$w = \frac{2 * cutoff\_frequency}{sampling\_frequency}$$

Where the $samplingfrequency$ is 102.4 Hz while the $cutofffrequency$ is 20 Hz, as previously mentioned. An example of the effect of this filter on the acceleration magnitude is reported in Figure 3.3.

Gait segment detection

As previously mentioned, a *gait cycle* is the sequence of events that occur during the walking process between two consecutive heel strikes of the same foot. Henceforth,

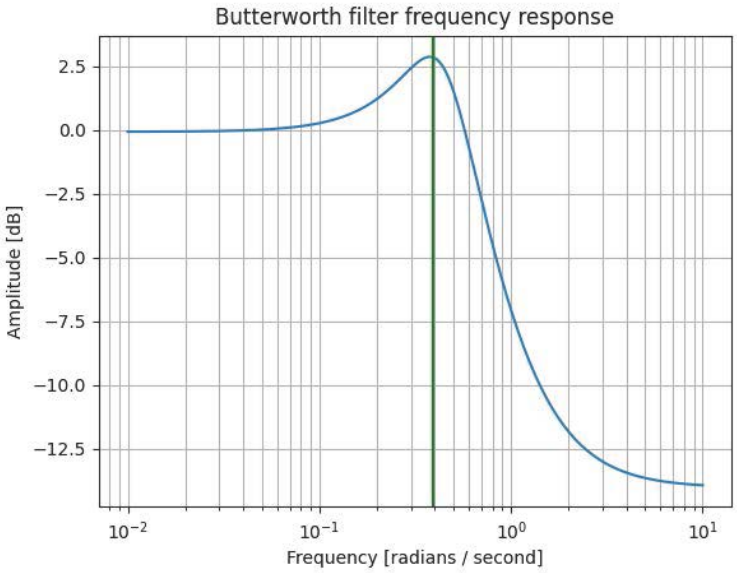


Figure 3.2: Frequency response of the Butterworth second-order filter used for signal filtering.

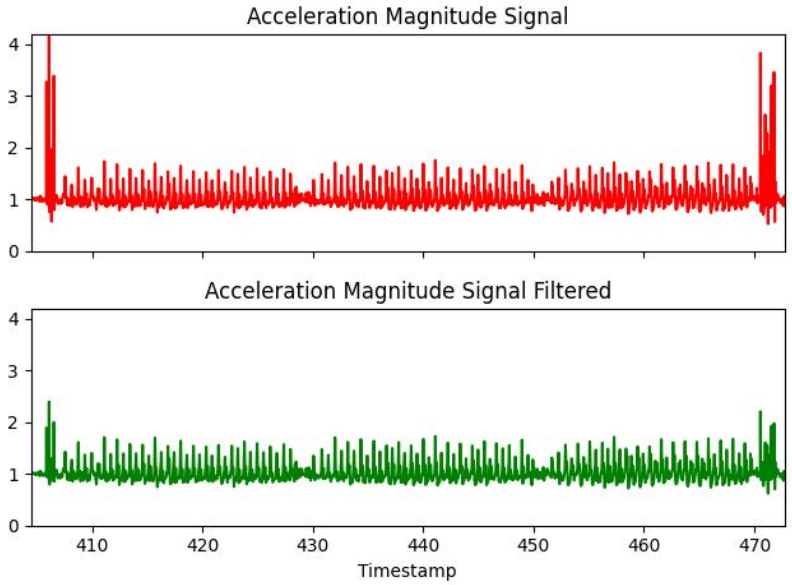


Figure 3.3: A comparison between the acceleration magnitude signal of a subject before and after the filter applications.

we will use the term *gait segment* to indicate four consecutive gait cycles. In this phase, gait segments are automatically identified by means of the walking detection algorithm described in (Cola et al., 2014), which is based on the analysis of the acceleration magnitude signal. The number of gait cycles per gait segment (four)

was determined empirically. According to our experiments, the selected number of cycles is able to capture the most important information on the subject's gait pattern. At the same time, a gait segment composed of just four gait cycles will enable the detection of such segments in home environments, where spatial characteristics may make long walks unlikely.

The gait segment detection algorithm performs the following:

- Peak detection stage
- Gait step detection stage
- Gait segment extraction stage

Peak detection: the filtered acceleration magnitude signal is scanned for peak detection. After filtering, the acceleration magnitude signal is scanned to detect peaks. In particular, a sample is considered a peak if its value is greater than the value of the next and the previous sample, but also it must be greater than the *peakth* threshold. *Peakth* expresses the peak's acceleration magnitude minimum values(Algorithm 1). Figure 3.4 depicts an example of a detected peak. The following algorithm presents the peak detection stage in detail.

Algorithm 1: Subject's peaks detection procedure

Data: Let $A = \{a_1, a_2, ..., a_n\}$ be the set of subject's acceleration magnitude, timestamp-sorted
Result: $P = \{p_1, p_2, ..., p_n\}$ set of subject's detected peaks
$P \leftarrow \{\}$;
foreach $a_i \in A$ **do**
  if $a_i > a_{i-1}$ & $a_i > a_{i+1}$ & $a_i > peakth$ **then**
    $P \leftarrow P \,||\, a_i$;

Gait step detection: a single gait step produces more than one peak on the acceleration magnitude signal trace. The *step detection* stage aims to group peaks that are temporally very close to each other, associating each group to a single step. Here, two additional thresholds are used:

- *gmaxint*, which expresses the maximum interval of time between two consecutive peaks that are related to the same step;
- *gdur*, that expresses the maximum duration of a group of peaks.

The algorithm works as follows: for each previously detected peak, the system groups it with the subsequent ones if the time distance is less than *gmaxint*; furthermore, the algorithm checks if the time interval between the first and the last peaks of

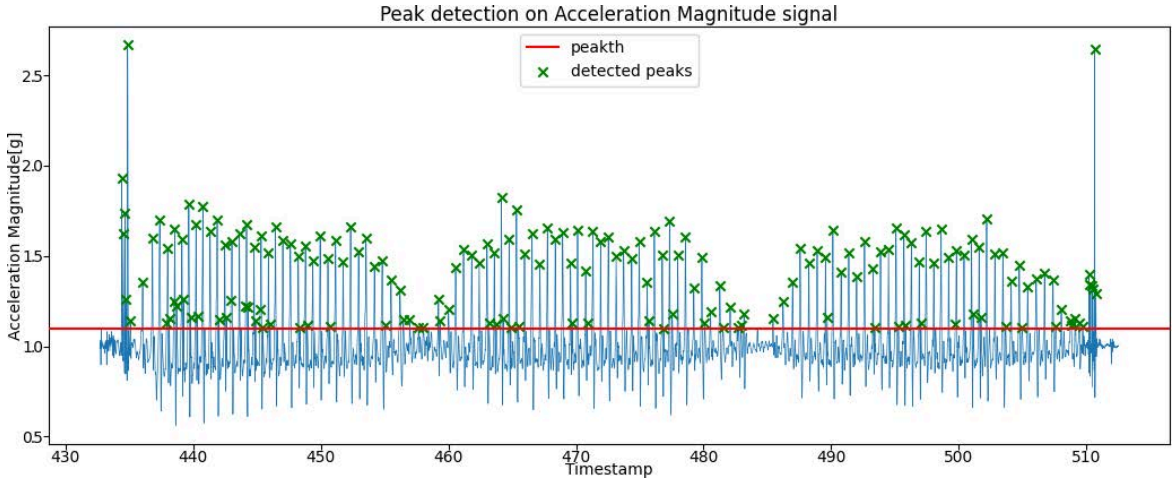


Figure 3.4: Example of the peaks identified by the gait segment detection algorithm.

the group is less than *gdur*. Hence, a *gait step* is detected by averaging the measures of all the peaks belonging to that particular group. A detailed description of the *gait step detection* algorithm is presented below.

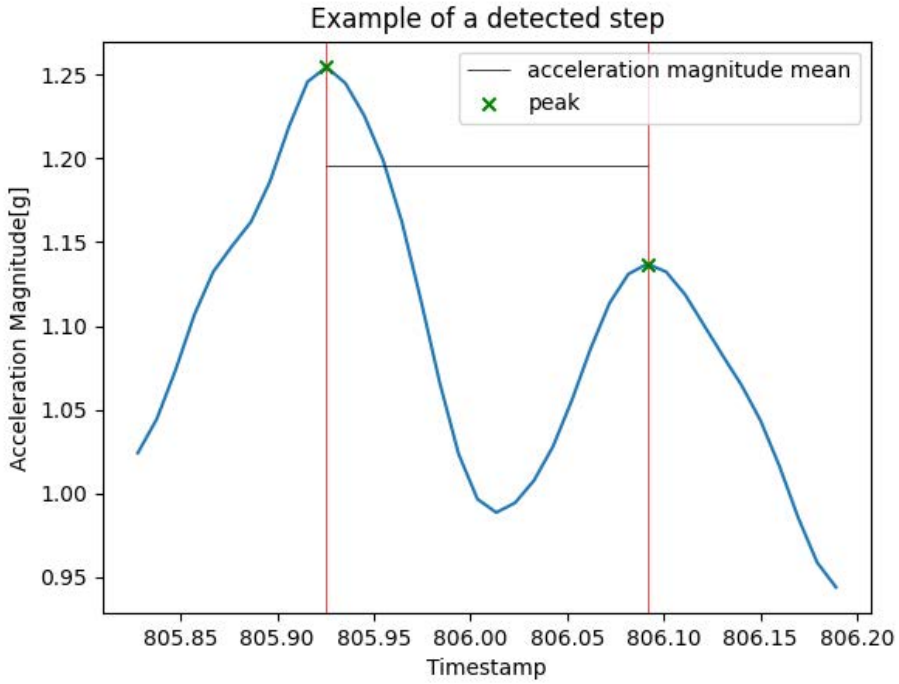


Figure 3.5: Gait segment detection: an example of a detected step.

An example of a detected step can be observed in Figure 3.5. Here, the vertical

Algorithm 2: Gait step detection algorithm

Data: Let $P = \{p_1, p_2, ..., p_n\}$ be the set of subject's acceleration magnitude peak instances sorted by timestamp
Result: $S = \{s_1, s_2, ..., s_n\}$ set of subject's detected steps
$S \leftarrow \{\}$;
Let $G = \{\}$ the detected groups of peaks;
foreach $p_i \in P$ **do**
  $starting\_peak \leftarrow p_i$;
  Let $T = \{p_i\}$ be the peaks' group under analysis;
  Let D_G the G duration;
  Let $P_S = \{p_{i+1}, p_{i+2}, ...\}$ the set of consecutive peaks;
  foreach $p_j \in P_S$ **do**
    if $D_G > gdur$ ***or*** $p_j - p_{j-1} > gmaxint$ **then**
      break;
    $T \leftarrow T \,||\, p_j$;
  $G \leftarrow G \,||\, T$;
foreach $g_i \in G$ **do**
  Let $P = \{p_0, p_1, ..., p_n\}$ be the set of peaks in g_i;
  Let T_M the mean of the instants of peaks $\in P$;
  Let A_M the mean of the acceleration magnitude of peaks $\in P$;
  $step\ [time] \leftarrow T_M$;
  $step\ [acceleration] \leftarrow A_M$;
  $S \leftarrow S \,||\, step$;

bars represent a peak group's start and end points.

Gait segment extraction: all the previously detected gait steps become the input of the *gait segment extraction* stage. This stage aims to group consecutive gait steps to form a gait segment. The number of gait steps grouped in one gait segment is expressed by the *gstepmin* parameter. Moreover, an additional threshold is exploited in order to express the maximum interval of time between two consecutive steps in the same gait segment. For each detected gait segment, a time window between the timestamps associated with the first and the last gait steps is extracted.

A detailed version of the *gait segment extraction* stage is presented below.

The action performed by the gait segment detection algorithm is independent of the orientation of the wearable device. However, the signals extracted from sensors positioned in different parts of the body could differ in shape or amplitude. For this reason, it is necessary to adapt the setup operation of the algorithm to the specific body position by tuning the thresholds used for peak-detection.

In the proposed method, the thresholds have been set as showed in Table 3.1.

To prevent anomalous segments from affecting the classifiers' performance, a

Algorithm 3: Subject's segments detection algorithm

Data: $S = \{s_1, s_2, ..., s_n\}$ set of subject's detected steps sorted by timestamp
Result: Let $SP = \{sp_1, sp_2, ..., sp_n\}$ and $EP = \{ep_1, ep_2, ..., ep_n\}$ be respectively the set of starting and ending points of the subject's windows extracted

```
SP ← {};
EP ← {};
step_index ← 0;
Let N be the number of the subjects detected steps;
while step_index < N do
    Let T = {} be the set of segments under analysis;
    Let D_T be the duration of the segment T;
    for i ← step_index to step_index + gstepmin by 1 do
        T ← T || s_i;
    if D_T < stepdmax then
        Let SS and ES respectively the first and last step contained in T;
        SP ← SP || SS[time];
        EP ← EP || ES[time];
        step_index ← step_index + gstepmin;
    else
        step_index ← step_index + 1;
```

Name	Value	Meaning
peakth	1.1 g	Peak Acceleration Magnitude minimum value
stepdmin	0.4 s	Step minimum duration
stepdmax	0.9 s	Step maximum duration
gdur	0.4 s	Peak group maximum duration
gmaxint	0.25 s	Max interval of time between two consecutive peak groups
gstepmin	8 s	Number of gait steps included in one gait segment

Table 3.1: Summary of the thresholds used by the gait detection algorithm.

segment filtering procedure has been implemented based on autocorrelation. More specifically, we compute unbiased autocorrelation coefficients (Moe-Nilssen and Helbostad, 2004a) on magnitude m samples as follows:

$$AC_k = \frac{1}{N-k} \sum_{i=1}^{N-k} r_i * r_{i+k},$$

where AC_k represents the k-th unbiased autocorrelation coefficient; k is the considered time lag; r_i represents the i-th magnitude m sample minus the average magni-

tude of the whole gait segment; N represents the number of samples in the whole gait segment.

Subsequently, a peak detection technique is used to find the first and the second dominant periods of the autocorrelation function (AC_DP1 and AC_DP2), which represent the estimated average duration of a step and average duration of a gait cycle, respectively (Moe-Nilssen and Helbostad, 2004b). The autocorrelation coefficient corresponding to the first dominant period (AC_C1) can be used to evaluate the regularity of consecutive steps, whereas the coefficient at the second dominant period (AC_C2) describes the regularity of consecutive gait cycles. The latter has been exploited to discard highly irregular gait segments. In particular, for each subject:

- the subject's gait segments are ordered in descending AC_C2 order;
- the first M segments, with the highest values of AC_C2, are selected for next phases.

From now on, let us set $M = 9$.

Frailty status assessment

The M gait segments that are not discarded in the previous phase, are used as input to *Feature Extraction*. Several features are computed in the time and time-frequency domains to capture the main characteristics of the subject's gait pattern. Features and feature selection are described and discussed in Section 3.2.

Frailty status is assessed in a two-stage process by means of a machine learning model. Let us define *gait instance* the vector of features extracted from a gait segment. First, each of the subject's gait instances is classified as belonging to the NR or R class. Then, the subject is classified according to the majority voting scheme shown in Algorithm 4.

3.2 The method's validation criteria

In this section, we describe the validation procedure for our method. In particular, we illustrate the experimental protocol used for building the ground-truth dataset, the techniques for selecting the set of features to be inputted to the predictive model, and how the performance of frailty status assessment is evaluated.

Participants and Experiment protocol

Adults aged 70+ years, consecutively accessing the Geriatrics outpatients clinic at Careggi academic hospital as patients or patient's caregivers, were assessed by geri-

Algorithm 4: Frailty status assessment procedure

Result: Subject's frailty status (R or NR)
Let $G = \{g_1, \ldots, g_M\}$ be the set of subject's gait instances;
foreach $g_i \in G$ **do**
  g_i is classified as R or NR;
end
Let $T = \{t_1, \ldots, t_M\}$, of length $M = 9$, be the set of labels assigned by the classifier;
Let T_{NR} be the total number of gait segments labeled as NR;
Let k_{NR} be the non-robust parameter, and let k_{NR} be equal to $0,40$;
$threshold \leftarrow \lceil M \cdot k_{NR} \rceil$;
if $T_{NR} \geq threshold$ **then**
  Subject is classified as NR;
else
  Subject is classified as R;
end

atricians. After exclusion of participants who reported physical dependence in at least one of Katz's basic activities of daily living (BADL) (Katz et al., 1963), as well as of those with conditions causing overt abnormalities of gait (stroke, Parkinson's disease, severe hip or knee osteoarthritis), 34 eligible subjects were enrolled.

The sample included 13 females (80.15 ± 6.80 years old, height 1.58 ± 0.04 m, weight 65.19 ± 11.80 kg) and 21 males (80.05 ± 6.30 years old, height 1.72 ± 0.07 m, weight 75.43 ± 8.91 kg). All the participants were screened for the presence of frailty based on Fried's criteria (Fried et al., 2001), in order to enroll older subjects in each of the three categories of *robust*, *pre-frail* and *frail*. As a result, the sample contained 23 NR (non-rubust, including 8 frail and 15 pre-frail) subjects and 11 R (robust) subjects.

After clinical evaluation, participants were asked to walk 60 meters at their preferred pace along a 20 meters long path, under the supervision of geriatricians. In order to study the relevance of the sensor's position, during the test the subjects wore two Shimmer3 devices, one placed on the lumbar back, and the other worn like a watch on the wrist (Figure 3.6). Shimmer3 is a wearable sensor embedding a tri-axial accelerometer (Shimmer, 2018), which was used to collect acceleration samples at $102.4Hz$. Sensors were positioned in the same way for all subjects, so that the directions of the sensor reference system with respect to the subject were consistent throughout experiments. Nevertheless, in the case of use in an uncontrolled environment, simple calibration procedures can be implemented for both considered on-body positions.

A signed, written consent to participate in the study was obtained from all participants. The study was conducted in accordance with the ethical principles of the

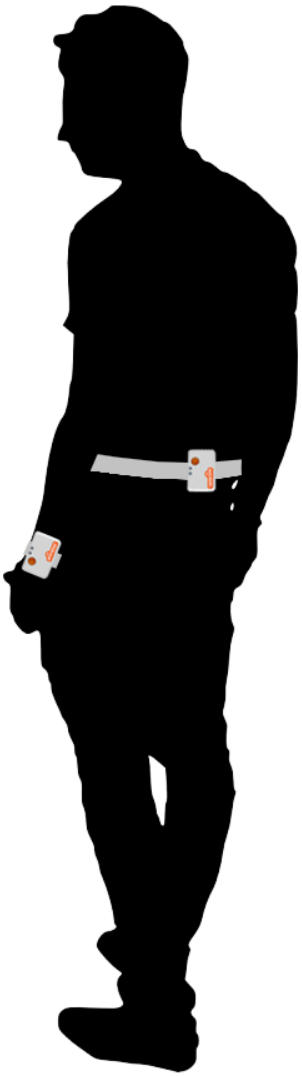


Figure 3.6: On-body sensors setup for frailty status assessment.

Declaration of Helsinki. Identifiable information was removed from the collected data to ensure participant anonymity. Ethical approval for this study was obtained from *Comitato Etico Regionale per la Sperimentazione Clinica della Regione Toscana* (approval n. 14834_oss of May 7, 2019).

Feature extraction

The set includes common statistical parameters used in signal processing (mean, median, standard deviation, minimum and maximum values, interquartile range (IQR), mean absolute deviation (MAD), root mean square (RMS), kurtosis, skewness and zero-crossing rate (ZCR)), calculated on acceleration components or Wavelet coefficients.

Table 3.2: List of extracted features in the time domain for frailty status assessment.

Feature	Components
Mean	x, y, z, m
Median	x, y, z, m
Standard deviation	x, y, z, m
Minimum value	x, y, z, m
Maximum value	x, y, z, m
Interquartile range	x, y, z, m
Kurtosis	x, y, z, m
Zero Crossing Rate	x, y, z, m
Mean Absolute Deviation	x, y, z, m
Root Mean Square	x, y, z, m
Average Absolute Variation	x, y, z, m
AC_C1	m
AC_C2	m
AC_DP1	m
AC_DP2	m
Cadence	-
Duration	-

Table 3.3: List of extracted features in the time-frequency domain for frailty status assessment.

Feature	Components
Mean	CWT coefficients
Median	CWT coefficients
Standard deviation	CWT coefficients
Minimum value	CWT coefficients
Maximum value	CWT coefficients
Interquartile range	CWT coefficients
Kurtosis	CWT coefficients
Skewness	CWT coefficients

In addition, we considered some other features previously used in gait analysis and fall detection studies: the *cadence*, defined as the ratio between duration of the

Table 3.4: Python Scikit-learn implementation of tested classifiers.

Model	**Parameters**
RandomForestClassifier()	n_estimators=50
GaussianNB()	default parameters
LogisticRegression()	penalty='l1', solver='liblinear'
MLPClassifier()	solver='lbfgs', activation='relu', hidden_layer_sizes=(12)
SVC()	kernel='poly', degree=2

gait segment and the number of performed steps, and the average absolute acceleration variation (AAV), which is computed on consecutive acceleration samples (Cola et al., 2014, 2017).

Features bringing information in the time-frequency domain were extracted by applying the Wavelet analysis. In particular, we used the *Continuous Wavelet Transform* (CWT) - see Chapter 2.5 - on the acceleration magnitude signal to study variations of power within gait segments. CWT is a signal processing technique used to extract time and frequency components, by combining translation and scaling of a special function called mother wavelet. The time-frequency representation of the gait segment enables the possibility to study the evolution of power during the movements related to gait.

Tables 3.2 and 3.3 list the full set of features considered in the feature extraction phase, for time and time-frequency domains, respectively.

Feature selection and frailty status assessment

Frailty status is assessed in a two-stage process by means of a machine learning model. In the first stage, gait instances are classified as belonging to a NR or R class. In the second stage, the subject is classified according to a majority voting scheme such as the one described in Algorithm 4. In order to maximise the performance of the classifier, we evaluated five different machine learning models: Random Forest, Gaussian naive Bayes, Logistic Regression, Multilayer Perceptron, and Support Vector Machine. All machine learning models have been implemented by means of the Python module Scikit-learn (Pedregosa et al., 2011). Details are shown in Table 3.4. The models' parameters have been chosen empirically. If the parameters are not given in the table, the default values have been used.

Models were tested by means of a Leave-One-Subject-Out cross-validation procedure: at each iteration, the gait instances of one subject were used as testing set, while the gait instances of other subjects were used as training set to build a clas-

sification model. A *feature selection* step was performed within the cross-validation procedure. At each Leave-One-Subject-Out iteration, features were selected using only the training set, so that the model was built without any test set information. This led to a different feature set at each iteration, according to the subjects belonging to the training set. The trained model was then used to classify the instances of the left-out subject as NR or R. Finally, subject classification as NR or R was based on the majority voting scheme. This procedure was repeated, each time leaving out a different subject as the testing set.

As far as *feature selection* is concerned, we used the One Way ANalysis Of VAriance (ANOVA) test between groups of features and the target. Specifically, for each feature, the values extracted from the R and NR subjects in the training set were compared, and the ANOVA F-statistic was computed. Features were then sorted in descending order by F-statistic, and the best k features were selected. In this case, $k = 25$.

As stated in the *Introduction*, one of the aims of this chapter is to show that walk related features in the time-frequency domain, computed by means of Wavelet analysis, can improve the accuracy of frailty status assessment. More precisely, we wanted to verify if the variation of power over time, captured by CWT, could bring valuable information for frailty status assessment purposes. To do this, we compared the use of CWT-based features against a system based on frequency domain features. Features in the frequency domain have been extracted by means of Fast Fourier Transform (FFT). Specifically, we computed some statistics to describe the distribution of FFT coefficients. Besides, we added the *1st_dominant_freq* feature, which represents the frequency with the highest corresponding value in FFT power spectrum. The list of FFT-based features is shown in Table 3.5.

In summary, we performed our evaluation using (i) only time domain features, (ii) a combination of time and time-frequency domains (i.e., CWT-based) features, and (iii) a combination of time and frequency domains (i.e., FFT-based) features. Therefore, the results of (i), (ii) and (iii) were compared to test our hypothesis.

From now on, let us consider NR subjects as positive and R subjects as negative classification results. *Accuracy*, *sensitivity* and *specificity* have been used to evaluate and compare performance of the classification models. Accuracy measures the ratio between the number of correctly classified NR and R subjects and the total number of subjects. Sensitivity, or true positive rate, measures the proportion of NR subjects correctly identified. Specificity, also called true negative rate, measures the proportion of R subjects correctly identified. In addition, *Receiver Operating Characteristic* (ROC) *Area Under the Curve* (AUC) has been calculated for every tested model.

Finally, let us recall that all the operations described so far were performed independently using only information extracted from the lower-back or the wrist sensor, hereafter called *LOWER BACK* and *WRIST* approaches, respectively.

Table 3.5: List of extracted features in the frequency domain.

Feature	Components
Mean	FFT coefficients
Median	FFT coefficients
Standard deviation	FFT coefficients
Minimum value	FFT coefficients
Maximum value	FFT coefficients
Interquartile range	FFT coefficients
Kurtosis	FFT coefficients
Skewness	FFT coefficients
First dominant frequency	-

3.3 Results and Discussion

In this section we report on the experiments we carried out following the validation procedure described in Section 3.2.

Performance of frailty status assessment

Data collected during the campaign (Section 3.2) was used for the Leave-One-Subject-Out cross-validation procedure of the machine-learning models (Section 3.2). Acceleration signals from both the wrist and the lower-back sensors were processed to detect gait segments (Section 3.1), from which a subset of twenty five features from the one listed in Section 3.2 were extracted, in order to create gait instances. The frailty status assessment scores obtained by the five chosen classifiers are summarized in Tables 3.6 and 3.7, for both the *LOWER BACK* and the *WRIST* approaches, respectively.

In the *WRIST* approach, for all three feature sets, Gaussian NB was able to classify subjects with good classification accuracy. The best scores were achieved using the combination of time domain and CWT-based features, with a ROC AUC of 0.87, 91.3% sensitivity (21 correctly classified NR participants out of 23) and 81.8% specificity (9 correctly classified R participants out of 11). Similarly, Random Forest achieved valuable results, with an AUC of 0.80, 95.6% sensitivity and 63.6% specificity.

As far as the *LOWER BACK* approach, Gaussian NB achieved the best performance using features in time and time-frequency domains, with an AUC of 0.75, 87.0% sensitivity (20 correctly classified NR subjects out of 23) and 63.6% specificity (7 correctly classified R subjects out of 11).

Table 3.6: Average results of frailty status assessment using a LOWER BACK sensor.

Features	Model	Acc.	Sens.	Spec.	AUC
TIME DOMAIN	**Gaussian NB**	0.76	0.87	0.55	0.71
	Random Forest	0.68	0.83	0.36	0.59
	Log. Regression	0.59	0.70	0.36	0.53
	ML Perceptron	0.65	0.70	0.55	0.62
	SVM	0.56	0.61	0.45	0.53
TIME DOMAIN + CWT-BASED	**Gaussian NB**	0.79	0.87	0.64	0.75
	Random Forest	0.76	0.87	0.55	0.71
	Log. Regression	0.74	087	0.45	0.66
	ML Perceptron	0.71	0.83	0.45	0.64
	SVM	0.74	0.83	0.55	0.69
TIME DOMAIN + FFT-BASED	**Gaussian NB**	0.76	0.87	0.55	0.71
	Random Forest	0.71	0.83	0.45	0.64
	Log. Regression	0.74	0.83	0.55	0.69
	ML Perceptron	0.50	0.35	0.82	0.58
	SVM	0.68	0.70	0.64	0.67

Table 3.7: Average results of frailty status assessment using a WRIST sensor.

Features	Model	Acc.	Sens.	Spec.	AUC
TIME DOMAIN	**Gaussian NB**	0.82	0.91	0.64	0.77
	Random Forest	0.74	0.78	0.64	0.71
	Log. Regression	0.76	0.87	0.55	0.71
	ML Perceptron	0.71	0.78	0.55	0.66
	SVM	0.74	0.78	0.64	0.71
TIME DOMAIN + CWT-BASED	**Gaussian NB**	0.88	0.91	0.82	0.87
	Random Forest	0.85	0.96	0.64	0.80
	Log. Regression	0.79	0.91	0.55	0.73
	ML Perceptron	0.76	0.87	0.55	0.71
	SVM	0.68	0.74	0.55	0.64
TIME DOMAIN + FFT-BASED	**Gaussian NB**	0.82	0.91	0.64	0.77
	Random Forest	0.76	0.91	0.45	0.68
	Log. Regression	0.71	0.83	0.45	0.64
	ML Perceptron	0.53	0.48	0.64	0.56
	SVM	0.74	0.78	0.64	0.71

From the results, it turns out that the wrist-worn device matched or outperformed the lower back-worn device for almost all the considered machine learning models. Therefore, according to our experiments, we argue that the signal extracted

from the wrist-worn device is most effective for discriminating between NR and R subjects. Such a result can be explained by considering the arm swing involved in human walking, to which a wrist-worn device is more exposed. In fact, due to the subject's better stability, the energy produced by wider arm oscillations and wrist rotations of a R subject is likely to be greater than that produced by a NR subject.

This information pattern can be observed in Figure 3.7, which depicts the boxplots generated using RMS values. The RMS, computed as the area under the acceleration amplitude signal, represents the energy of the signal. RMS values of gait segments belonging to R subjects are generally higher than values computed for NR subjects. Even though such a difference is observable in both the *WRIST* and the *LOWER BACK* approaches, in the former the distinction is more evident. To statistically confirm the significance of RMS, we performed a one-way ANOVA test. More precisely, RMS was used as the dependent variable and the frailty category (R or NR) as the independent variable. The results are shown in Table 3.8. For both *WRIST* and *LOWER BACK*, RMS appears statistically significant in distinguishing R and NR subjects, with a $p - value < 0.001$ (the $p < 0.005$ criterion was used to test for statistical significance).

According to these findings, it appears that arm swing brings valuable information in frailty status assessment. As these gait-related features can be better captured by a wrist-worn device, we can conclude that our method implementation is suitable to be embedded in a common smartwatch, thus enabling continuous assessment of frailty without requiring the adoption of additional devices.

Table 3.8: Results obtained in ANOVA statistical test, for RMS in *WRIST* and *LOWER BACK*.

	R	**NR**	$p - value$
WRIST	1.085 ± 0.291	1.047 ± 0.114	< 0.001
LOWER BACK	1.057 ± 0.037	1.043 ± 0.063	< 0.001

Wavelet Analysis

As mentioned in Section 3.2, models were trained and tested using (i) only time domain features, (ii) time domain + CWT-based features, and (iii) time domain + FFT-based features. This was done to investigate whether the use of acceleration features in the time-frequency domain may improve the performance of predictive models. As hypothesized, it appears that the CWT-based approach improves the classification task sensibly, independently of the position chosen for the sensor. Gaussian NB proved to be the best model in frailty status assessment, for all three considered feature sets. In the case of training performed by means of time domain + CWT-based

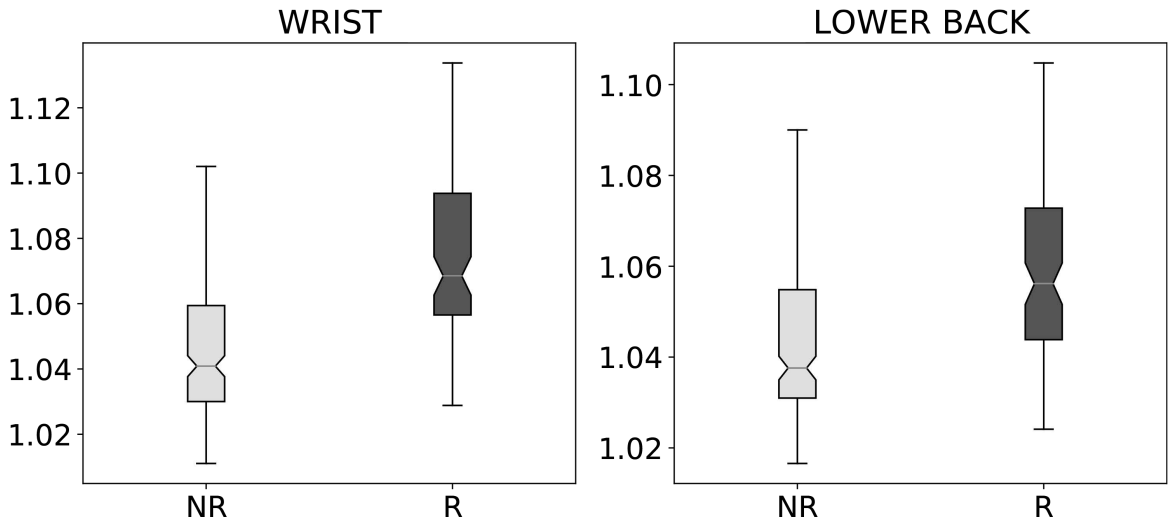


Figure 3.7: Box plot representation of the Root Mean Square computed on the acceleration magnitude of R vs NR signals, in both WRIST and LOWER BACK.

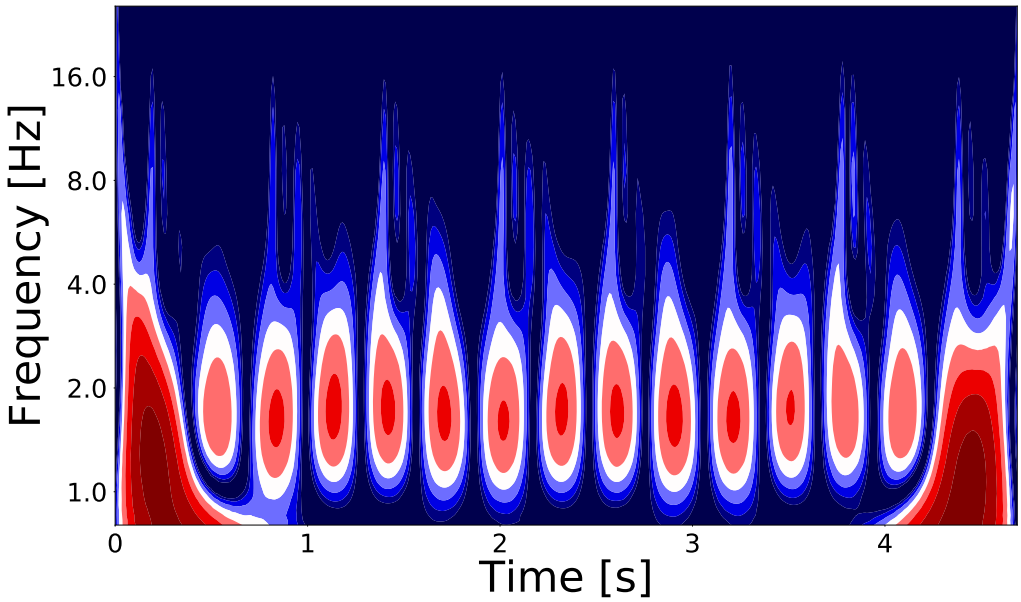
wrist-derived features, however, an important step forward was made, with a 0.1 increase in the AUC score.

It is worth mentioning that all the features described in the previous sections are always extracted from gait segments consisting of four gait cycles. This leads to gait segments composed of a variable number of samples, since the length of a gait segment depends on the subject's cadence. Performing frequency-based features extraction on gait segments of variable length introduces differences into the outputs of the frequency analysis, on which the frequency-based features are computed. Nevertheless, in the current study, these differences concern the frequency band $[0, 0.26]Hz$, which from an in-depth analysis appears to contain nonsignificant information for frailty status assessment. We have chosen not to report here the results of the frequency bands significance analysis, as it is beyond the scope of this study.

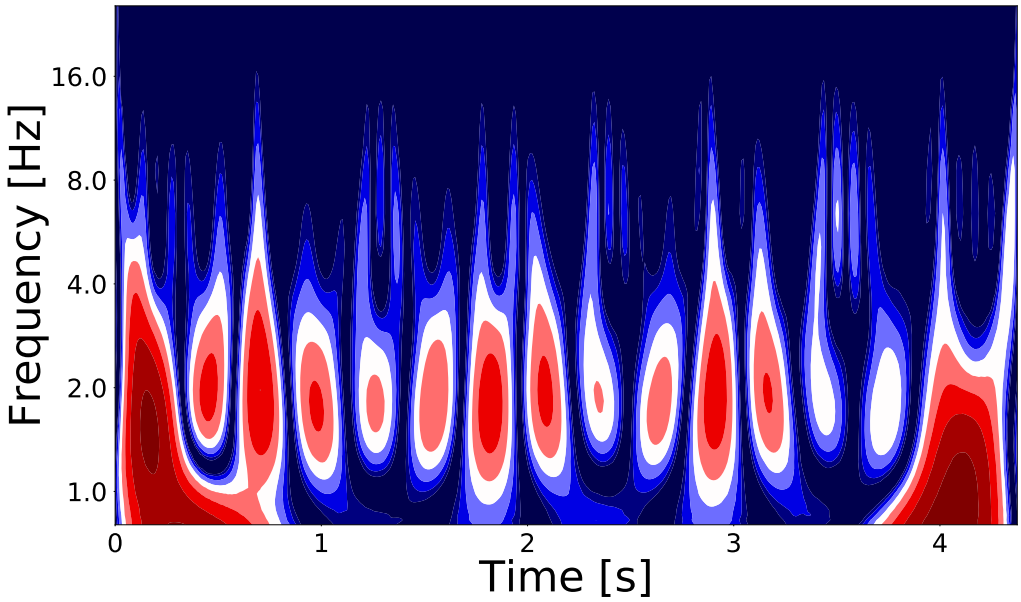
Now, let us discuss why Continuous Wavelet Transform enhanced the characterization of gait segments in view of the machine-learning classification of frailty status. In particular, we argue that CWT applied to gait segments can better capture the nature of oscillations produced by different cyclic phases involved in a subject's walk, such as arm motions, torso rotation and heel strike.

Let us first recall that, while the FFT only brings information about the frequency components of a signal, the CWT enables the study of signal power distribution in the time-frequency domain.

It worth recalling that variations of power over time can be visualized by means of scalograms (such as the ones shown in Figure 3.8), where colors represent the in-

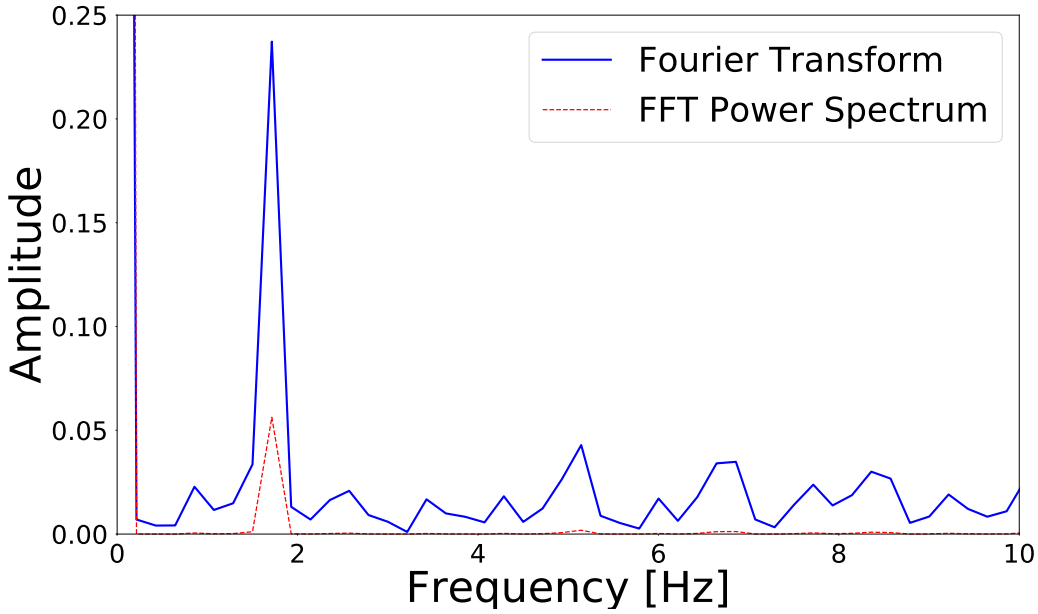


(a) R subject

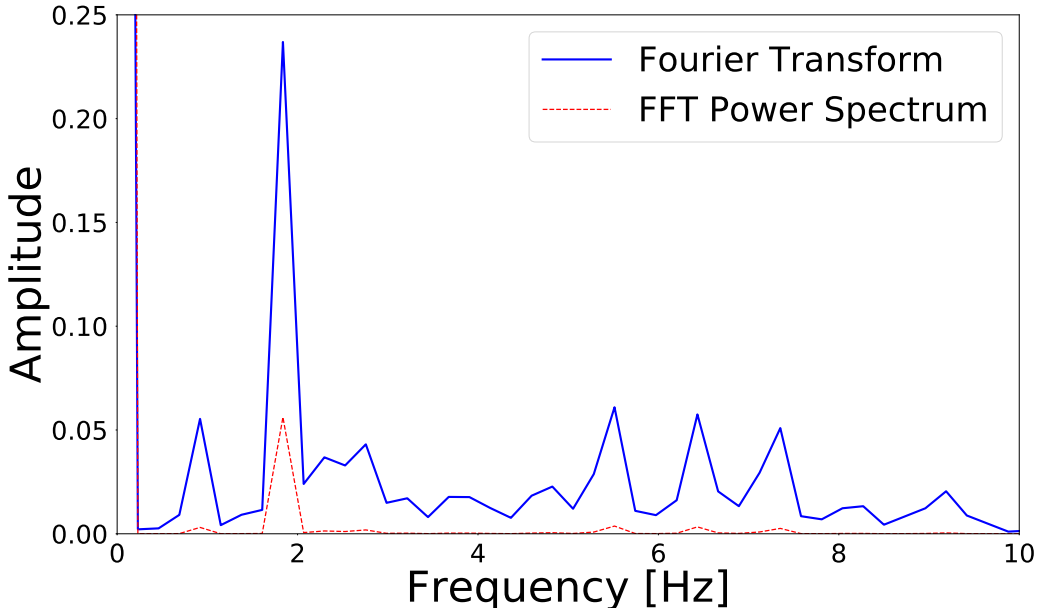


(b) NR subject

Figure 3.8: Power spectrum of the Continuous Wavelet Transform in robust and non-robust subjects.



(a) R subject



(b) NR subject

Figure 3.9: Power spectrum of the Fast Fourier Transform in robust and non-robust subjects.

tensity of power - from blue (low power) to red (high power). In Figures 3.8 and 3.9 we report the outputs of CWT and FFT analysis applied to two sample gait segments, in the case of a R and a NR subject, respectively. Signals were recorded by the wrist sensor (similar considerations also apply to the lower back-worn sensor). Again, the gait segments used in this comparison are composed of 4 gait cycles. Here, red areas of scalograms correspond to higher levels of power released during stronger oscillations in gait, within a given frequency range (y-axis values), over a particular time interval (x-axis values). Notably, areas recurring with similar shape can be attributed to a regular gait. In other words, they characterise a walk-related pattern that produces regular oscillations over time. Large red areas at the boundaries of scalograms are the result of truncation of the gait segments and do not bring any useful information.

As can be seen in Figure 3.8a, most of the signal power released during the gait of a R subject is within the $1.5 - 3Hz$ frequency range, in accordance with the corresponding peak of the Fourier transform shown in Figure 3.9a. Notably, the scalogram evidences a certain regularity in the of power released during the gait. The same cannot be said of the gait signal produced by a NR subject, whose scalogram is shown in Figure 3.8b. Here, even though red areas still depict a gait activity with a similar level of associated power, it is clear that power is not released with the same regularity. It should be noted that such a difference is not evident when comparing the FFT pictures (e.g., Figures 3.9a vs. 3.9b).

These results suggest that a higher level of power distributed regularly along time is associated with better stability during gait. In contrast, an irregular power distribution over time reflects a high gait variability, whose correlation with frailty has already been explored in previous studies. Montero-Odasso et al. demonstrated that a high gait variability is a marker of the loss of complexity in the dynamics of the gait pattern, and it is associated with frailty status (Montero-Odasso et al., 2011). In our experiments, we found that a higher level of power correlates with a more emphasized arm swing, which is also known to be positively related to "global gait stability" (Bruijn et al., 2010). These findings are also in line with those of Mirelman et al., who reported that aging is associated with decreased arm swing amplitude (Mirelman et al., 2015).

Chapter 4

Wavelet-based analysis of gait for automated frailty assessment

In this Chapter we aim to statistically analyze the output of Continuous Wavelet Transform to improve the assessment of frailty based on gait-related parameters derived from a wrist-worn device. The chapter is organized as follows. First, Section 4.1 presents the proposed method. In Section 4.2, the the experimental design is presented, together with the validation procedure. Finally, in Section 4.3 the results achieved are presented and discussed.

4.1 Method

Figure 4.1 shows a schematic view of the proposed approach. A wearable device positioned on the subject's wrist is used for signal acquisition. After a preprocessing step aimed at reducing noise, a gait segment detection technique identifies gait segments on the whole signal trace, which are then used as input for CWT. Finally, band-based feature extraction is responsible for gathering information from the CWT output, which is used to determine which *Frequency Bands* (FBs) contain the most valuable information for frailty status assessment.

In the following we provide more detail on the techniques used throughout the process. The experimental setup and the statistical and machine learning evaluations are discussed in Section 4.2.

Data acquisition and preprocessing

Acceleration is sampled at 102.4 Hz. All the acceleration components (**x**, **y**, **z** axes and acceleration magnitude **m**) are converted into g units. A second order Butterworth low-pass filter with a cut-off frequency of 20 Hz was applied to the signal, as body movements typically have frequency components below 20 Hz (Antonsson

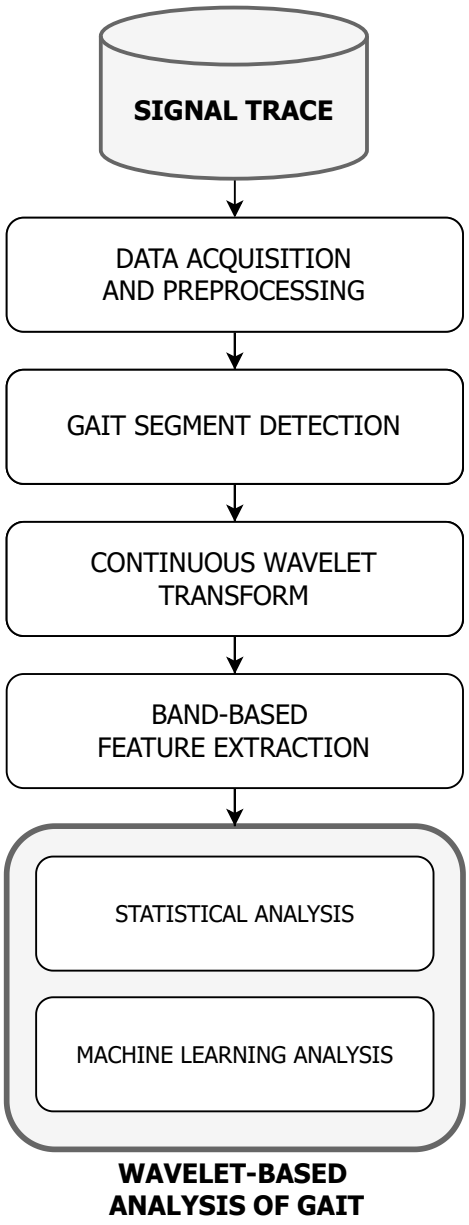


Figure 4.1: Flowchart of the proposed method for wavelet-based analysis of gait.

and Mann, 1985). Filtered acceleration components are then provided to the gait segment detection module.

Gait segment detection

A gait cycle is the sequence of events that occur during the walking process between two consecutive heel strikes of the same foot (Avvenuti et al., 2018). Henceforth, we will use the term *gait segment* to indicate four consecutive gait cycles. In this

phase, gait segments are automatically identified by means of the walking detection algorithm described in (Cola et al., 2014), which is based on the analysis of the acceleration magnitude signal.

Continuous Wavelet Transform

Wavelet analysis is applied to gait segments detected in the previous step. In particular, we use the *Continuous Wavelet Transform* (CWT) on the acceleration magnitude signal to obtain a representation of the gait segment into the time-frequency domain, in order to study variations of power over time. CWT is a signal processing technique for analyzing a time series containing non-stationary power at different frequencies (Daubechies, 1990). More specifically, CWT allows us to analyze local variations of power, by decomposing a time series into different frequency components.

Starting with a Mother Wavelet (MW), the time-frequency representation of the time series is obtained by a combination of scaling and translation applied to MW. Let us call a the scale factor. Scaling a wavelet simply means stretching it (if $|a| > 1$) or compressing it (if $|a| < 1$). On the other hand, translating a wavelet means shifting its position in time. Hence, translation and scaling describe the position of the wavelet in time and frequency domains, respectively. In the proposed method, we choose Morlet wavelet as the MW, since it has been proven to be the a good choice for CWT applied to gait signals (Khandelwal and Wickström, 2016). It is worth mentioning that, knowing the MW, scale values can be easily converted to frequency values.

Band-based Feature Extraction

The phase discussed above produces as output a matrix M_{CWT} of coefficients c_{ij} for each gait segment, where i denotes the frequency value, obtained from the scale factor, and j denotes the instant in time. Usually, this matrix of coefficients is plotted through the *wavelet power spectrum* (or *scalogram*). The scalogram is the absolute value of CWT plotted as a function of frequency (on a logarithmic scale on the y-axis) and time (on the x-axis). From visual inspection, some FBs present similar characteristics among classes, whereas other FBs present distinctive patterns for each frailty class. As already mentioned, we aim to quantitatively identify the FBs that can best distinguish subjects according to their frailty status. To accomplish this, we split the matrix of coefficients into several sub-matrices M_{s-e}, where s and e represent the lower and upper bounds of the considered FB, respectively. For instance, M_{0-1} represents the sub-matrix of coefficients between 0 and 1 Hz. The studied FBs are shown in Table 4.1.

Table 4.1: List of studied frequency bands

Lower bound (Hz)	Upper bound (Hz)
0	0.5
0.5	1.5
1.5	2.5
2.5	3.5
3.5	6
6	10
10	20

For each M_{s-e} a set of features is extracted, to further describe the distribution of the wavelet coefficients in that particular FB. These includes common statistical measures, such as minimum and maximum values, mean, median, variance, interquartile range (IQR), coefficient of variation (CoV), median absolute deviation (MAD), and skewness. Moreover, for each sub-matrix a function is built, to summarize the trend of the FB wavelet coefficients in time. Henceforth, we will refer to this as *Wavelet Coefficients Evolution function* $WCEf$. Let us consider a sub-matrix M_{s-e}, defined in FB $[s, e]$ and in time interval $[t_0, t_{max}]$. $WCEf(t_j)$ is then:

$$WCEf(t_j) = \frac{\sum_{i=s}^{e} c_{ij}}{N}. \tag{4.1}$$

where t_j is the considered instant in time - the M_{s-e} column index j - and N represent the number of rows in M_{s-e}. In other words, for each $t_j \in [t_0, t_{max}]$ we computed the average of the coefficients in column j of M_{s-e}. Finally, root mean square (RMS) of $WCEf$ was computed as an additional feature. Let us define *band-based gait instance* the vector of band-based features extracted from a gait segment

4.2 Experimental protocol

Participants and Experimental protocol

A total of 34 adults (13 females: 80.15 ± 6.80 years old, height 1.58 ± 0.04 m, weight 65.19 ± 11.80 kg; 21 males: 80.05 ± 6.30 years old, height 1.72 ± 0.07 m, weight 75.43 ± 8.91 kg) were assessed by geriatricians. Geriatricians used Fried's criteria to screen participants for the presence of frailty (Fried et al., 2001), in order to enroll older subjects in each of the three categories of *robust*, *pre-frail* and *frail*. As a result, the sample contained 8 frail (F), 15 pre-frail (PF) and 11 robust (R) subjects. After clinical screening, participants were asked to walk 60 meters at their preferred pace,

Table 4.2: robust vs. non-robust classification scores

Frequency Band (Hz)	**Subject Classification Accuracy**
$[0, 0.5]$	0.65
$[0.5, 1.5]$	0.65
$[1.5, 2.5]$	0.77
$[2.5, 3.5]$	0.68
$[3.5, 6]$	0.68
$[6, 10]$	0.65
$[10, 20]$	0.70

under the supervision of the geriatric staff. During this experiment, subjects wore a Shimmer3 device on their wrist, like a watch. Shimmer3 is a wearable device embedding a tri-axial accelerometer (STMicro LSM303DLHC) (Shimmer, 2018), which was used to collect acceleration samples at $102.4Hz$.

Statistical analysis and Machine Learning classification

The importance of each FB for distinguishing between the three groups (F, PF and R) was determined using *One-Way ANalysis Of VAriance* (ANOVA). In addition, we performed *independent samples t-tests* for between-groups comparisons, in order to inspect the differences between pairs of frailty classes: R vs PF, PF vs F and R vs F. The $p < 0.005$ criterion was used to test for statistical significance.

To further compare CWT coefficients in different FBs, we trained and tested seven Machine Learning models, each time using only features extracted from one of the FB shown in Table 4.1. Specifically, we implemented seven Random Forest binary classifiers by means of the Python module Scikit-learn (Pedregosa et al., 2011), in order to distinguish robust subjects from non-robust ones (i.e., pre-frail or frail). We tested each model by performing Leave-One-subject-Out cross validation: at each iteration, band-based gait instances of 33 subjects were used as training set, while the band-based gait instances of the remaining subject formed the test set. Then, the subject was classified according to a majority voting scheme: if more than 40% of the band-based gait instances were classified as non-robust, the subject was classified as non-robust. Early identification of non-robust subjects (pre-frail and frail) may enable prompt clinical intervention to delay the evolution of frailty.

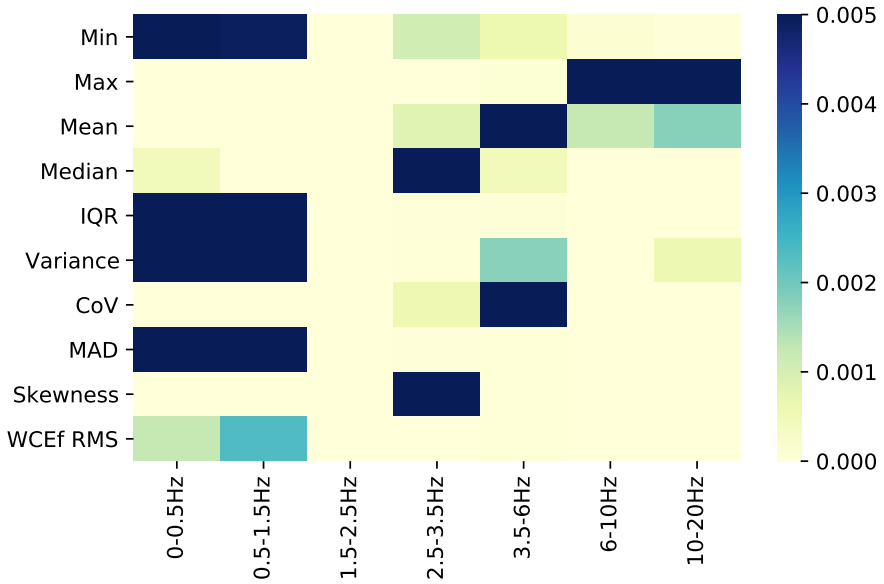


Figure 4.2: p-values computed in ANOVA: R vs PF vs F.

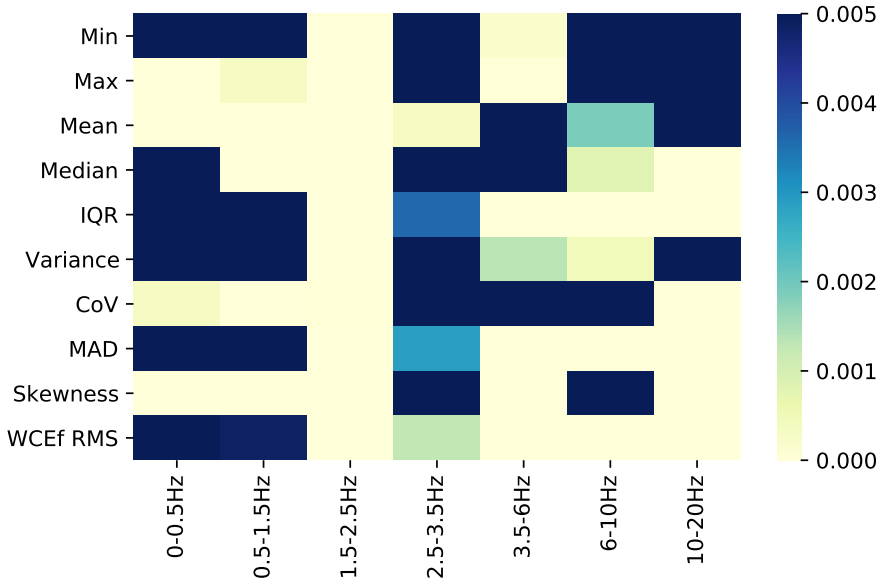


Figure 4.3: p-values computed in R vs PF t-test.

4.3 Results and Discussion

In this study, we hypothesized that specific CWT frequency bands may carry valuable information for gait analysis aimed at frailty assessment. On the other hand, some frequency bands may lower the performance in frailty assessment, and should

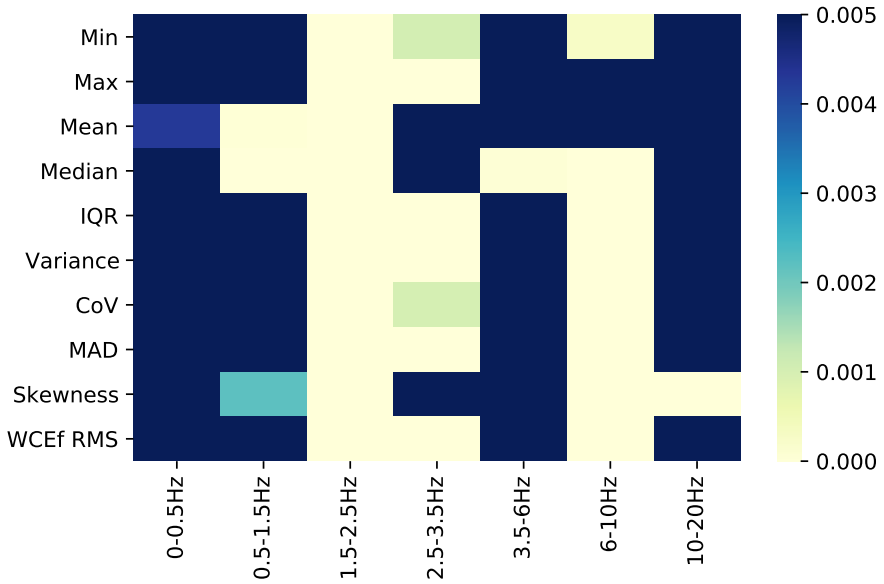


Figure 4.4: p-values computed in PF vs F t-test.

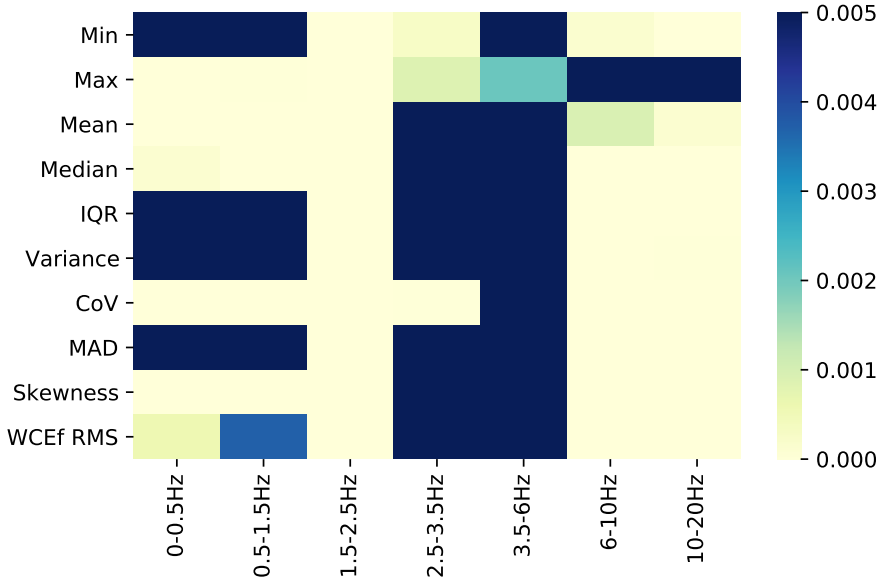


Figure 4.5: p-values computed in R vs F t-test.

not be considered. The heatmaps in Figures 4.2, 4.3, 4.4, and 4.5 show the $p-values$ computed in the statistical tests. Dark blue colored squares indicate statistically non-significant features. It turns out that the features extracted in the $[1.5, 2.5]Hz$ FB proved to be highly significant ($p < 0.001$) in distinguishing subjects belonging to the three frailty classes, both in ANOVA and t-tests. This frequency range is in line

with the typical step frequency of walking (Antonsson and Mann, 1985). Moreover, among considered features, RMS of $WCEf$ reported the lowest p, showing that the energy of the CWT coefficients evolution signal in the $[1.5, 2.5]Hz$ band could be extremely valuable for frailty assessment.

Finally, we used band-based extracted features to train different random forest classifiers, in order to further verify our hypothesis. The accuracy scores, obtained in testing these R vs. PF or F binary classifiers, are listed in Table 4.2. Here, the accuracy score refers to the number of correctly classified subjects, as indicated in Section 4.2. The classifier trained with only the gait instances in the $[1.5, 2.5]Hz$ FB achieves the best results, in line with the result obtained from statistical testing. This supports our initial hypothesis, as a specific FB of the CWT provides a significantly more valuable information for frailty assessment.

Chapter 5

A Deep Learning approach for frailty status assessment

In Chapters 3 and 4, we applied a series of signal processing and statistical tools to analyze a set of gaits extracted from 34 subjects. We then observed how the output of CWT can increase the performance of frailty status assessment, as the scalograms extracted from subjects belonging to the same frailty class seem to present similar patterns.

In this chapter, we aim to verify this hypothesis using a Deep Convolutional Neural Network to detect these patterns in the output of Continuous Wavelet Transform.

5.1 A brief introduction to Deep Convolutional Neural Networks

Deep Learning is the field of Machine Learning characterized by multiple processing layers - called *deep layers* - each extracting information from the previous layer and computing values for the subsequent one. One of the main advantages of Deep Learning is the ability to learn hidden patterns from the input and to apply this information to make predictions.

Our system uses a particular class of deep neural networks called *Convolutional Neural Networks* (CNN). A CNN is an interleaved set of feed-forward layers that implements convolutional filters. Each layer in the network represents a group of artificial neurons able to originate a high-level abstract feature (Ravì et al., 2016). Typically, a CNN includes the following layers:

Convolutional: in these layers, a *kernel* (i.e., a filter) is applied to detect patterns in the input. The kernel moves along the input, and convolution is computed. Different kernels detect different features. Finally, an activation function will use the convolution results to produce the layer output, called *feature map*. Figure 5.1 depicts

a graphic representation of the above description.

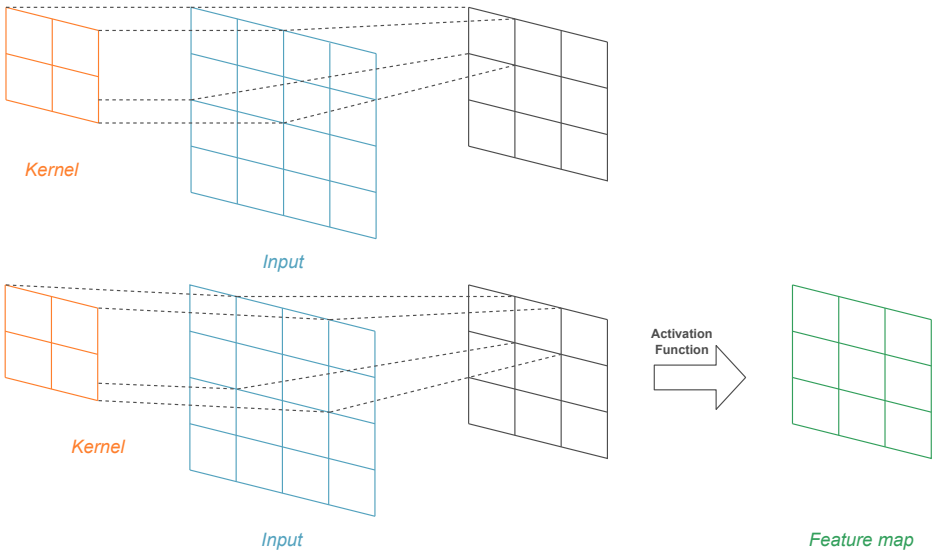


Figure 5.1: Example of a convolutional layer.

Pooling: convolutional layers generate feature maps sensitive to the features' position. A flipped input will produce a completely different feature map. CNNs use pooling layers to obtain a robust feature map. These layers use moving windows to summarize the information in the feature map through an aggregation function - e.g., max, mean, or min functions. Figure 5.2 shows an example of MaxPooling.

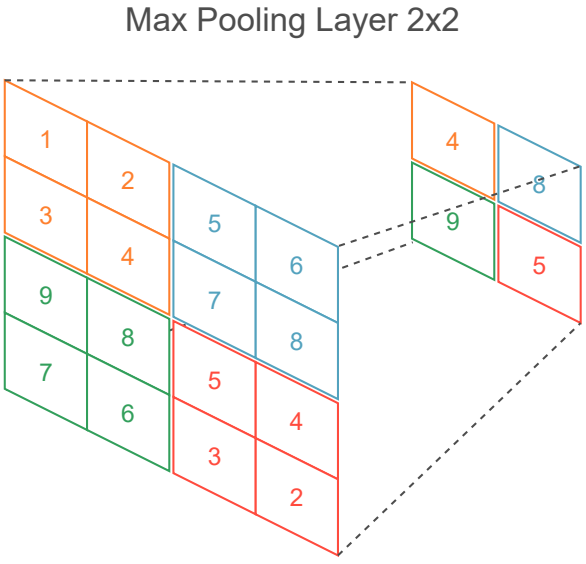


Figure 5.2: Example of a max-pooling layer.

Fully-connected (or **Dense**): each neuron is connected to all the neurons of the subsequent layer. These layers detect global features by aggregating previously extracted local features.

Finally, it is worth describing another mechanism used in CNNs: **Dropout**. Dropout is a regularization technique in which a group of randomly selected nodes is disconnected from the network at each iteration. In other words, incoming and outgoing connections are removed from the chosen neuron forcing the network to improve the learning techniques. Dropout decreases the probability of suffering the *overfitting* problem in the Neural Network. An overfitted network is characterized by high performance on the training set and lack of generalization capability, i.e., poor performance on unseen data.

5.2 Materials and methods

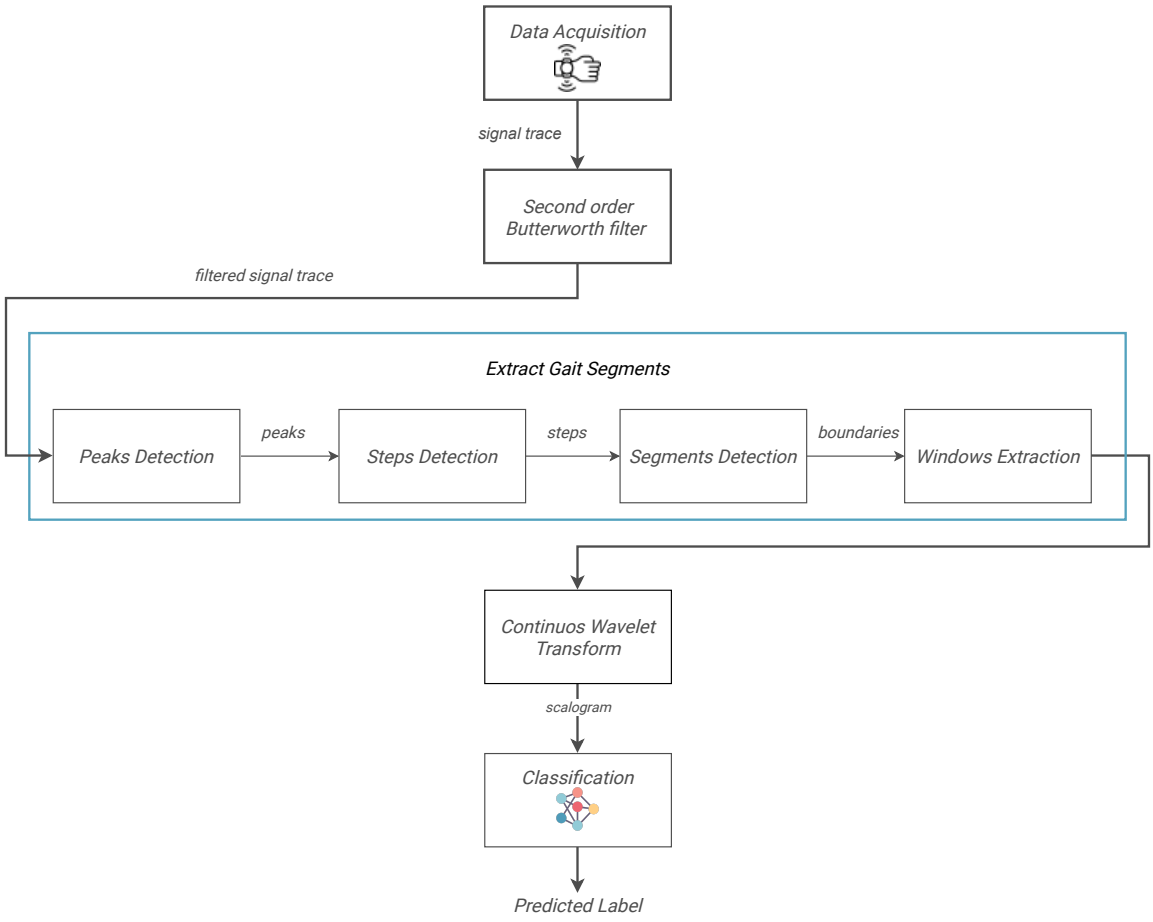


Figure 5.3: Flowchart of the CNN-based frailty status assessment proposed method.

This section describes the proposed method for a CNN-based frailty status assessment. Figure 5.3 depicts a summarized view of the method. The dataset used in this chapter has already been described in Chapter 3. The experimental protocol, the worn device, the signal preprocessing techniques, and the gait detection algorithm are the same. The only difference is the number of gait steps included in a gait segment. A peculiarity of the training phase of CNNs is the need for a large dataset to increase classification performance without running into problems such as overfitting. Due to the limited dimension of the dataset, we decided to extract gait segments composed of one gait cycle instead of four (two consecutive gait steps instead of 8). In this way, we increased the dataset size by four times.

A data summary is presented in Table 5.1. For the reader's convenience, we recall that the sample contained 23 NR (non-robust, including 8 frail and 15 pre-frail) subjects and 11 R (robust) subjects. This binary classification still allows geriatricians to identify the individuals who may need further clinical evaluation, namely pre-frail and frail. Indeed, early identification of the latter categories enables prompt clinical intervention, which may lead to better health outcomes for the subject.

Table 5.1: Description of the dataset used for CNN-based frailty status assessment.

	Females	Males
Number of subjects	13	21
Age	80.15 $\pm$ 6.80 years old	80.05$\pm$6.30 years old
Height	1.58 $\pm$ 0.04 m	1.72$\pm$0.07 m
Weight	65.19 $\pm$ 11.80 kg	75.43 $\pm$ 8.91 kg

Figure 5.3 shows the proposed method's flowchart.

In the following sections, we explore the signal transformation and the final classification process based on a Convolutional Neural Network.

Continuous Wavelet Transform

As previously mentioned, this chapter aims to verify the hypothesis that a Deep Convolutional Neural Network can detect frailty-related patterns in the scalograms obtained through Continuous Wavelet Transform. For this reason, the gait segments detected in the previous stage become the input to a CWT processing module.

A description of the CWT signal processing tool has already been presented in Section 2.5. Figure 5.4 shows an example of transformation performed in the current stage.

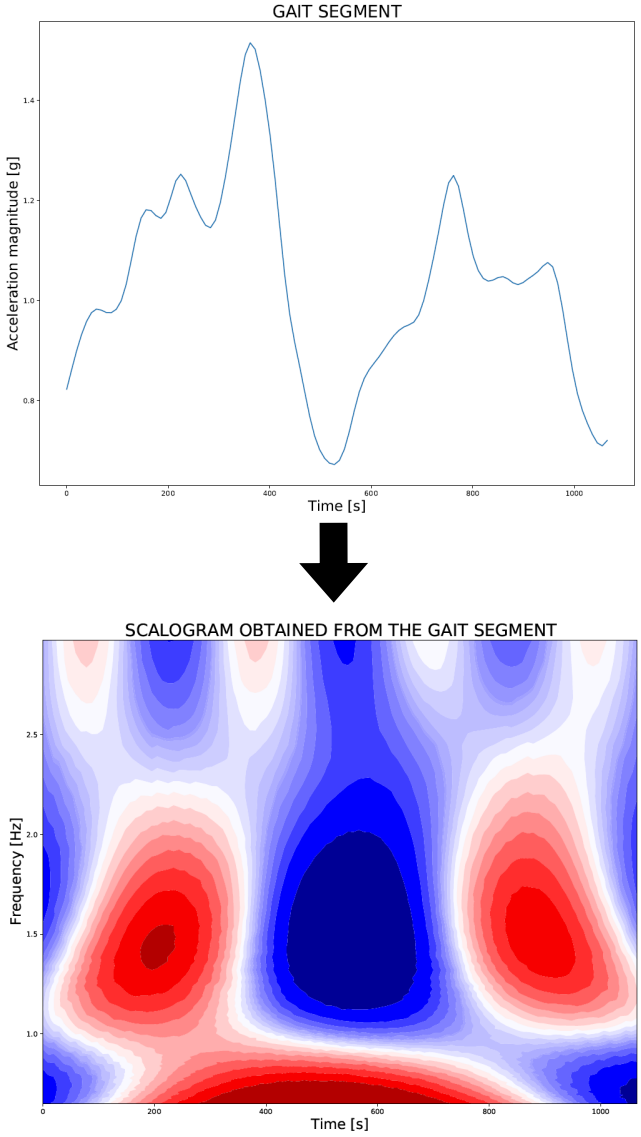


Figure 5.4: Example of transformation performed by Continuous Wavelet Transform.

Table 5.2: Dataset summary after CWT.

Frailty class	Total number of scalograms
Robust	1124
Non-robust	2643

The *base model*: CNN for Human Activity Recognition

Table 5.2 shows the number of scalograms available for training and testing the CNN. Notably, the dataset is not large enough to train high-performance learners

such as convolutional networks. Therefore, there is a need to perform part of the training phase using data from different domains. This methodology is referred to as transfer learning (Weiss et al., 2016).

There are very few publicly available Deep Learning models trained on scalograms. We thus decided to build a CNN for Human Activity Recognition (HAR) and trained it using a set of scalograms generated by CWT applied to publicly available wrist-derived accelerometric data. This way, the *base model* - i.e., the pre-trained model - was ready to extract features from scalograms for HAR aims. After, we re-trained the last layers of the *base model* using the dataset described in Table 5.2 to obtain our CNN-based frailty status classifier.

Activity	Windows
Walking	**2218**
Jogging	**2169**
Stairs	**2185**
Sitting	**2245**
Standing	2281
Typing	2162
Brushing Teeth	2200
Eating Soup	2207
Eating Chips	2214
Eating Pasta	2140
Drinking from Cup	2275
Eating Sandwich	2146
Kicking (Soccer Ball)	2208
Playing Catch w/Tennis Ball	2214
Dribbling (Basketball)	2242
Writing	2269
Clapping	2200
Folding Clothes	2227

Table 5.3: Summary of the WISDM Smartphone and Smartwatch Activity and Biometric Dataset.

Pre-training: Human Activity Recognition For this first phase, we used the WISDM Smartphone and Smartwatch Activity and Biometric Dataset (Weiss et al., 2019). It contains a high number of accelerometer-derived data extracted from a wrist-worn sensor. Each of the 51 subjects in the dataset performed 18 activities for 3 minutes each. During the experiment, each participant wore a smartwatch embedding an accelerometer. This sensor was used to extract activity-related signals at a sampling rate of $20Hz$.

A general schema of the operations performed to generate the training set of our *base model* is depicted in Figure 5.5. In particular, we split the acceleration magnitude signal into 96-sample windows - 4.8 seconds - and applied CWT to these windows. We tested different mother wavelets to understand the best one for our class of problems. A dataset summary is shown in Table 5.3. The activities in bold are the gait-related ones - i.e., the ones used to train the *base model*.

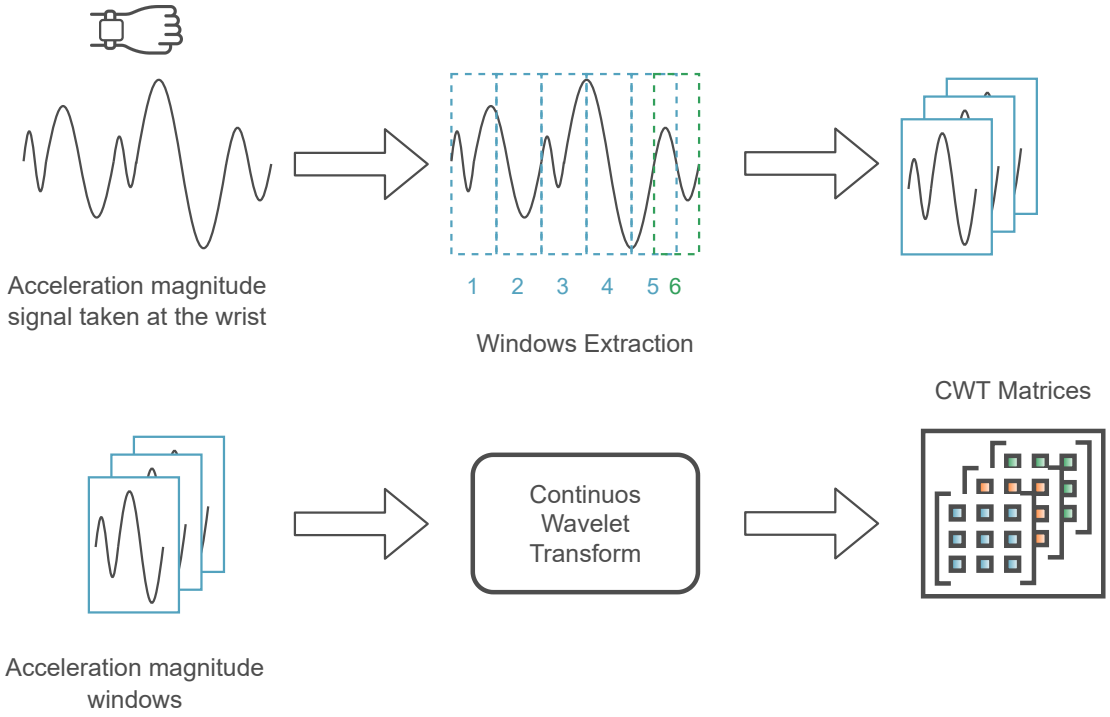


Figure 5.5: A schema of the preprocessing pipeline for the activity classification task

An example of the different Continuous Wavelet Transform obtained for the different activities and mother wavelets can be found in Section 2.5.

Finally, Figure 5.6 shows the architecture of the *base model*. We used the Rectified Linear Unit as the activation function in the hidden neurons. For the output layer - each one related to an activity - we chose the softmax activation function, which expresses the probability that the input belongs to that particular class.

CNN-based frailty status assessment

This chapter aims to use information obtained from CWT output to classify subjects according to their frailty status. After completing the training phase of the base model, we applied the CWT to the 34 subjects' signals described in Section 3.1. We used the resulting scalograms to re-train the base model's last layers. In particular,

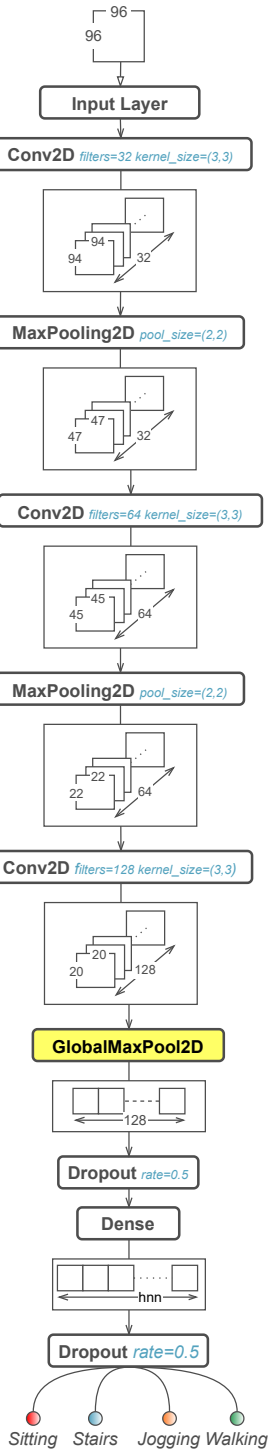


Figure 5.6: CNN for Human Activity Recognition: the architecture

Algorithm 5: Subject classification procedure

Data: Let $G = \{g_1, g_2, ..., g_n\}$ be the set of subject's scalograms
Result: Subject frailty status (R or NR)
Let $P = \{\}$ the set of model's prediction on the scalograms;
foreach $g_i \in G$ **do**
    Let L be the predicted label of g_i (R or NR);
    $P \leftarrow P \,||\, g_i$;
Let T be the threshold -empirically selected- on the percentage of segments classified as NR;
Let P_R the number of scalograms classified as R;
Let P_{NR} the number of scalograms classified as NR;
$PP \leftarrow P_{NR}/(P_R + P_{NR})$;
if $PP \geq T$ **then**
    Subject is classified as NR ;
else
    Subject is classified as R;

we trained a binary classifier to distinguish between robust and non-robust subjects. First, the NR or R class is determined for each one of the subject's scalograms. The subject is then categorized using a majority voting scheme, as indicated in Algorithm 5.

Figure 5.7 shows the re-trained layers.

5.3 Experiments

Leave One Subject Out Cross Validation

In order to validate our model, we applied a Leave One Subject Out Cross Validation (LOSO CV), which also represents a good choice when dealing with datasets of signals acquired from different subjects. LOSO CV has already been described in Section 3.2. At the beginning of each iteration, the base model is loaded and artificial neurons are initialized with random weights. The model is then trained and validated with a dataset composed of the windows extracted from $n - 1$ subjects. Finally, the windows of the left-out subject are given as input to test the model, and the corresponding predictions are stored along with the ground truth label. At the end of the last iteration, the system computes the model's score by comparing the ground truth and the predicted class for each of the scalograms present in the dataset.

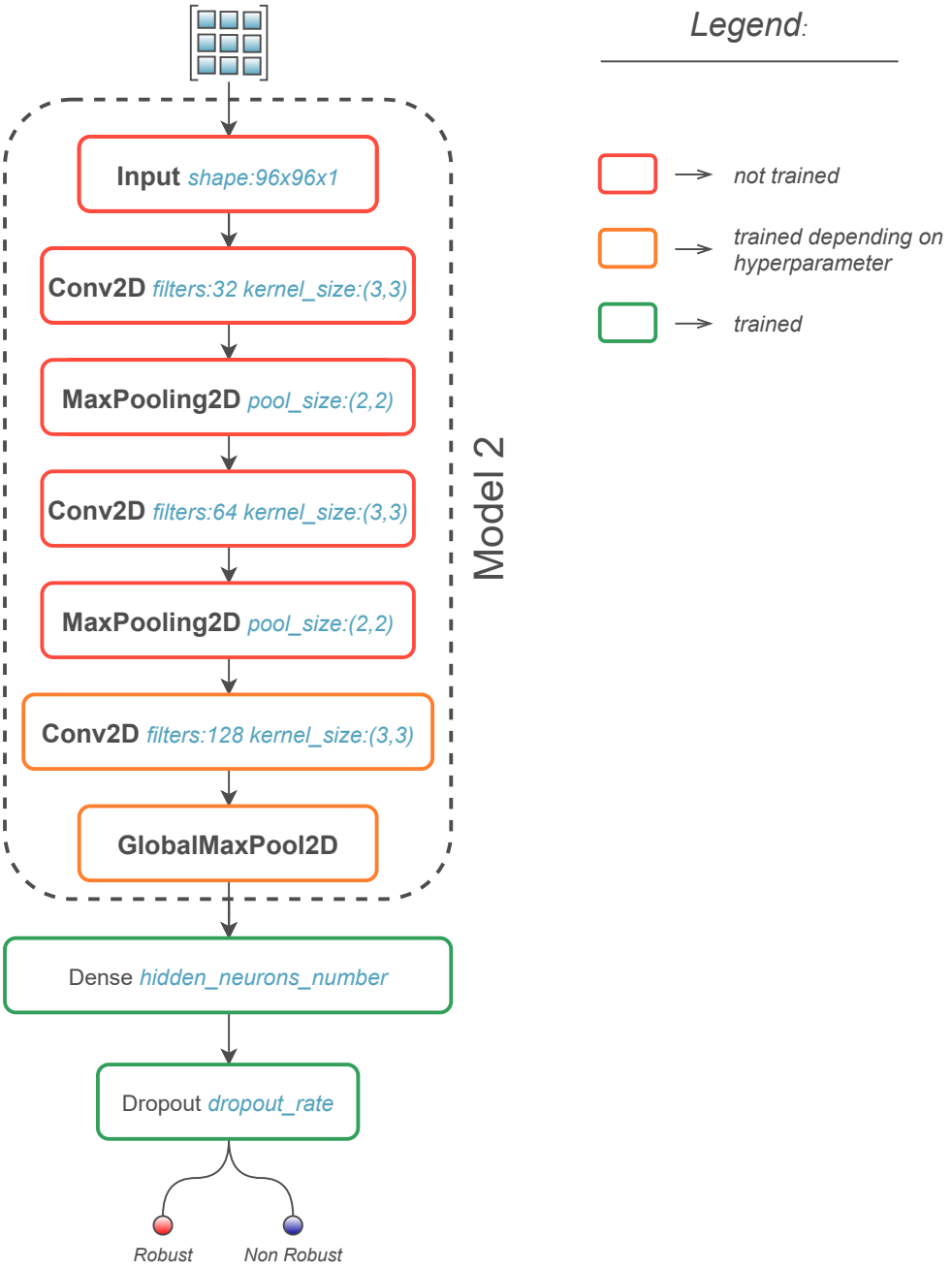


Figure 5.7: Application of Transfer learning for frailty status assessment.

Metrics

From now on, let us consider NR subjects as positive and R subjects as negative classification results. *Accuracy*, *sensitivity* and *specificity* have been used to evaluate and compare performance of the classification models. Accuracy measures the ratio between the number of correctly classified NR and R subjects and the total number of subjects. Sensitivity, or true positive rate, measures the proportion of NR subjects correctly identified. In the medical field, sensitivity is a crucial metric since classifying as negative a subject that indeed is positive can seriously impact the subject's health. Specificity, also called true negative rate, measures the proportion of R subjects correctly identified. In addition, *f-1 score* has been calculated for every tested model.

Tensorflow

In our experiments, we used the Tensorflow API Abadi et al. (2015) for Python. Briefly, Tensorflow is a popular library developed by Google for developing Machine Learning and Deep Learning models. Particularly interesting for further improvements, Tensorflow allows exporting models compatible with smartphones, smartwatches, and constrained settings, enabling the possibility to build mobile applications. In particular, for developing the Deep Learning models proposed in this thesis, the Keras Chollet et al. (2015) library - included in the Tensorflow API - was used to enable the high-level development of deep neural networks.

5.4 Results

In this section, we present the results obtained with the *base model* in the HAR task and the results obtained in the frailty assessment task.

Base model results Figure 5.8 shows the accuracy score achieved in HAR for the different mother wavelets, using different numbers of neurons in the hidden dense layer.

The *morl* mother wavelet is the one that performed the best in most of the cases, achieving very good results in this first phase. For this reason, we decided to use this model for the second phase of the proposed method, which consists of re-training using the frailty-labeled dataset. No critical changes are achieved by changing the number of neurons. Thus, we decided to use 2^5 neurons.

Frailty status assessment results Table 5.4 presents the results of the CNN-based frailty classifier in distinguishing NR scalograms from R ones, while Table 5.5 shows the results for subject classification.

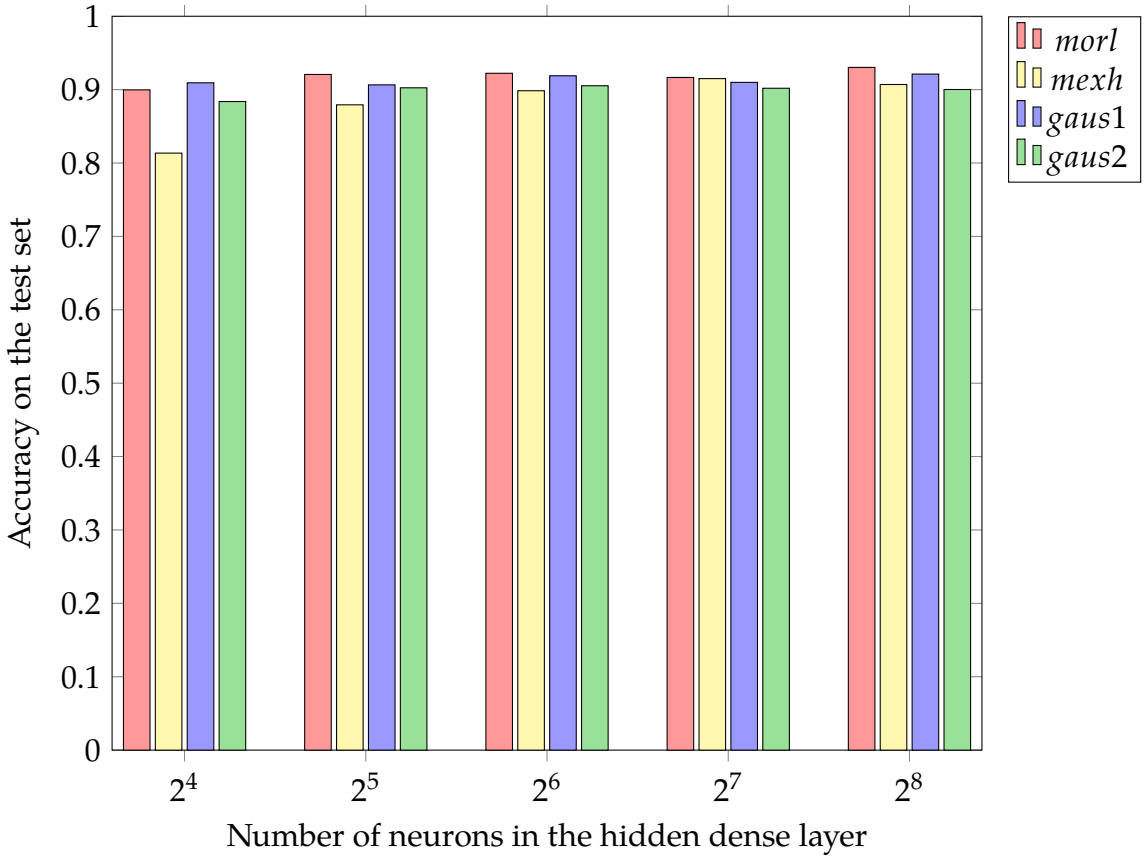


Figure 5.8: *Base model* accuracy score.

Table 5.4: Scalogram-wise results obtained for CNN-based frailty status assessment.

Mother Wavelet	hyperparameters	f1 score	Accuracy	Sensitivity	Specificity
gaus2	base_lr=0.001 dropout=0.25 hnn=128 epochs=10 unlocked	0.838	0.756	0.896	0.429
gaus1	base_lr=0.075 dropout=0.25 hnn=256 epochs=5 unlocked	0.817	0.756	0.908	0.261
mexh	base_lr=0.001 dropout=0.25 hnn=1024 epochs=7 unlocked	0.806	0.712	0.853	0.382

Table 5.5: Subject-wise results obtained for CNN-based frailty status assessment.

Mother Wavelet	hyperparameters	f1 score	Accuracy	Sensitivity	Specificity
gaus2	base_lr=0.001 dropout=0.25 threshold=80% epochs=10 hnn=128 unlocked	0.842	0.800	0.800	0.800
gaus1	base_lr=0.075 dropout=0.25 threshold=80% epochs=5 hnn=128 unlocked	0.789	0.714	0.833	0.5
mexh	base_lr=0.001 dropout=0.25 threshold=40% epochs=7 hnn=1024 unlocked	0.800	0.700	0.947	0.273

In Table 5.6, the subjects incorrectly classified are reported. As we expected, the model misclassified pre-frail subjects 9, 22, and 26, since subjects belonging to this class seem to share patterns with both frail and robust subjects.

Subject ID	Prediction	Ground Truth	Frailty Class
9	R	NR	PF
15	NR	R	R
22	R	NR	PF
26	R	NR	PF
27	NR	R	R
29	R	NR	F

Table 5.6: Subjects improperly classified by the CNN-based frailty classifier.

The results obtained in the subjects classification task are very promising and suggest that CWT applied to gait analysis could play a crucial role in frailty detection systems. Although we used transfer learning to overcome the dataset size problem, we note that the system still suffers from this limitation. Systems based on deep neural networks need very large datasets to identify the distinguishing features of the classes involved in these studies.

In chapter 3, we applied five shallow Machine Learning models for the same classification task: RandomForest, Gaussian naive Bayes, Logistic Regression, Multilayer Perceptron, and Support Vector Machine. Gaussian naive Bayes is the one that achieved the best performance, being able to classify subjects with a classification accuracy of 88%, 91.3% sensitivity (21 correctly classified NR participants out of 23), and 81.8% specificity (9 correctly classified R participants out of 11).

Let's compare the performance of deep and shallow learning models. The score achieved by the latter is slightly better than the one obtained with Deep Learning. This difference can be found in the ratio between system complexity and dataset dimensions: in general, shallow learning performs better than Deep Learning with a dataset characterized by limited dimensions. This shows that Deep Learning models, although they are very complex and have produced great results on some problems, are not universally better than Shallow Learning ones. Deep Learning methods often require large amounts of training data and high computational cost. In many cases, simpler ML models, such as the ones we used in Chapter 3, actually perform better.

Chapter 6

Automated, ecologic assessment of frailty using a wrist-worn device

In this Chapter, we study and discuss the potential of unobtrusive, wearable devices in objectively assessing frailty through unsupervised monitoring in real-world settings. The chapter is organized as follows: Section 6.1 presents the method we suggested for automatic frailty assessment based on wearable accelerometers. The experiment's setup, design decisions, and validation are discussed in Section 6.2. Finally, the achieved results are presented in Sections 6.3.

6.1 The proposed method

Figure 6.1 depicts the method's flowchart. In a nutshell, 24-hour acceleration data are collected through a sensor embedded in a wrist-worn wearable device. After preprocessing, the acceleration trace is sent to a segmentation module and split into 10-second segments. These segments are labeled according to their content: gait, other motor activity, or rest. Other motor activities and gait segments are used in the evaluation of the subject activity level (SAL). Furthermore, a subset containing the gait segments with the highest level of signal energy is used for the feature extraction phase. Finally, the extracted features are fed into the training phase of a Machine Learning classifier for the assessment of the frailty status. Specifically, the SAL and the result of the Machine Learning classification are used to assess whether the subject is robust (R) or nonrobust (NR).

Data acquisition and preprocessing

The acceleration components (**x**, **y**, **z** axes) are collected through the wearable device at a sampling rate of 102.4 Hz and converted into g units. Afterwards, the acceleration magnitude **m** is computed by means of the formula $m = \sqrt{x^2 + y^2 + z^2}$. As

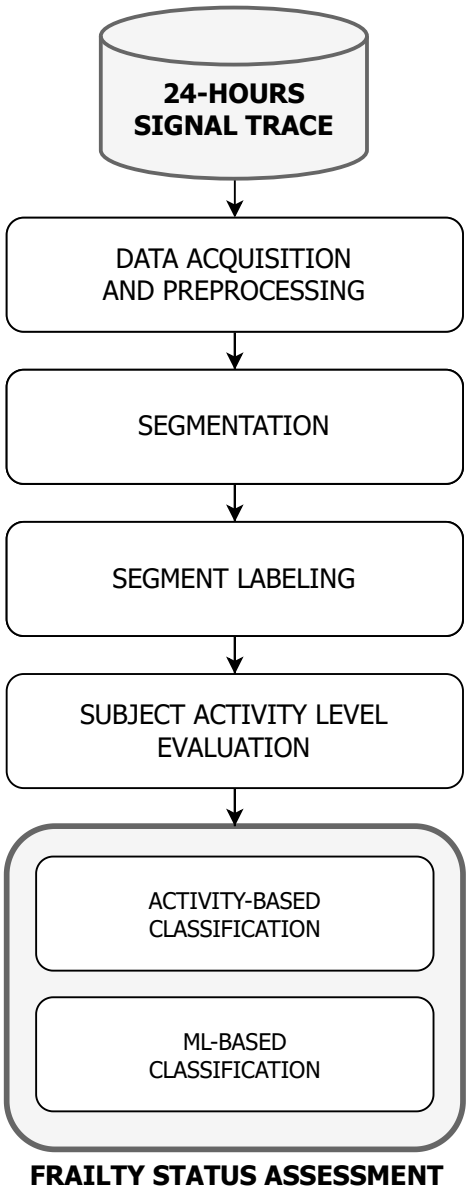


Figure 6.1: Flowchart of the proposed method for ecologic assessment of frailty using a wrist-worn device.

highlighted by the authors in (Antonsson and Mann, 1985), body movements typically have frequency components below 20 *Hz*. Hence, a second-order Butterworth low-pass filter with a cut-off frequency of 20 *Hz* is applied to all the acceleration components. The signal trace is now ready for the segmentation module.

Segmentation

In order to improve our system's generalization, we split the signal trace into segments of fixed duration. This allowed us to increase the dataset size and identify critical frailty-related features in small-sized segments. The duration was determined empirically. In detail, since the segments are given as input to a gait detection algorithm and the frailty status is mainly identified based on gait characteristics, the choice of duration has to respect the following requirements:

- even the shorter segments must allow capturing gait instances that accurately document the walking pattern of the subject and its regularity;
- given the limited space usually available in the home environment, walks are usually composed of a limited number of steps. Therefore, excessive duration of the segments might lead to incorporating other activities, together with gaits. This would result in a decline in the performance of frailty status classification.

According to our experiments, the best trade-off to optimize the classification would be $duration = 10s$. From now on, we will refer to the output of the segmentation phase with *10s segments*.

Segment labeling

Several researchers have stated the importance of human activity levels in assessing different clinical conditions. In this chapter, we evaluate the subjects' activeness by exploring the nature of their 10s segments. In particular, we apply a two-stage algorithm to analyze the content of each 10s segment and discriminate them among gait, other motor activity, and rest. Other motor activities and gait segments are then used to evaluate the subject's activity level. More details are presented and discussed in Section 6.1.

In the first stage, gait segments are automatically identified through the gait detection algorithm developed in (Cola et al., 2014) and adapted in (Minici et al., 2021). The algorithm is based on the analysis of the acceleration magnitude signal to detect gait cycles in the segment, where a gait cycle is the sequence of events that occur during two consecutive heel strikes of the same foot. A segment is labeled as *gait segment* if it contains at least four consecutive gait cycles. Otherwise, all those segments that remain unlabeled would be moved to the next stage.

In the second stage of the segment labeling algorithm, other motor activities are identified by means of the *standard deviation* of the acceleration magnitude. A segment is labeled as *other motor activity* if the standard deviation value is higher than a threshold determined experimentally *threshold_movement*. All the remaining segments are labeled as *rest segments* and discarded from the system.

Subject Activity Level

Objectively measured physical activity and sedentary behavior are associated with frailty in community-dwelling older adults. Prior understanding of this relationship relied on self-reported, subjective measures of physical activity, which are often biased (Tolley et al., 2021). In our study, we introduce the Subject Activity Level (SAL) parameter to assess this relationship in an objective fashion. In particular, we aim to use this information, related to the whole day (24 hours), along with the gait characteristics, which relate to a very limited time interval instead.

After labeling all the segments identified in a subject, we then compute the percentage of daytime spent walking, performing other motor activities, or resting. Let GS, $OMAS$, and RS be the number of 10s segments labeled as *gait segments*, *other motor activity segments*, and *rest segments*, respectively. Let TS be the total number of 10s segments identified in a subject. We will have:

$$GL = \frac{GS}{TS}$$

$$OMAL = \frac{OMAS}{TS}$$

$$RL = \frac{RS}{TS}$$

where GL, $OMAL$, and RL are the subject's gait level, other motor activities level, and rest level, respectively. It is worth noting that $GL + OMAL + RL = 1$.

The SAL is calculated as:

$$SAL = (1 - \alpha)OMAL + \alpha GL, \tag{6.1}$$

where α is a parameter to weight GL and $OMAL$. In particular, we hypothesized that the importance of GL and $OMAL$ can vary with different experimental setups and diseases. Therefore, the introduction of the alpha parameter allows us to study the activity that most influences a particular clinical condition in order to maximize the performance of the evaluation system. The study of alpha in our frailty status assessment system is provided in Section 6.2.

Frailty status assessment

In this phase, the SAL and gait segments are used to determine the frailty status of the subject. The design of the frailty status assessment algorithm is depicted in Figure 6.2. Frailty status is assigned in a two-stage process. In the first stage, the class assignment is based only on SAL; thus, we refer to *activity-based classification*. When SAL alone cannot provide a reliable prediction of the subject's class, classification is

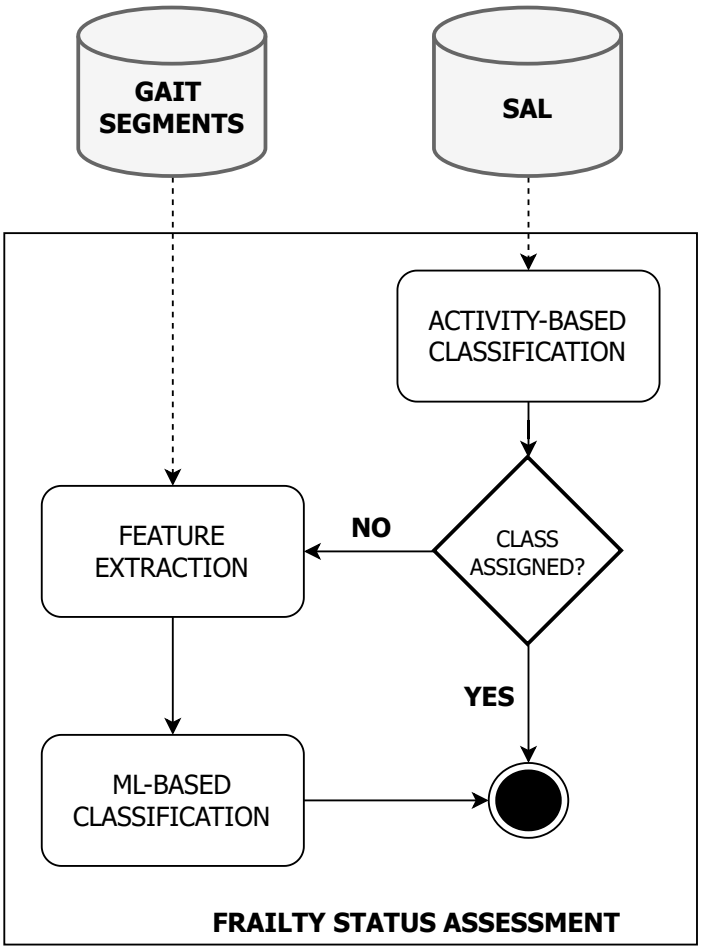


Figure 6.2: Two-stage frailty status assessment.

based on the second stage, which relies on a Machine Learning (ML) model: *ML-based classification*.

In the first stage based on SAL, we assume that a subject with a high SAL is more likely to belong to the R class, as frailty is associated with low levels of activity throughout the day. Hence, a subject with a low SAL value exhibits sedentary behavior and is more likely to belong to NR. Let AL_R and AL_{NR} be the threshold values to assign the R and NR classes to a subject, respectively.

$$\begin{cases} SAL \leq AL_{NR} \rightarrow subject\_class = NR \\ SAL \geq AL_R \rightarrow subject\_class = R \\ otherwise \rightarrow ML - based\ classification \end{cases}$$

More details on threshold estimation are presented in Section 6.2.

The ML-based classification module is a slightly modified version of what we already proposed in (Minici et al., 2021). The walks performed by subjects during the

day might have very different patterns, and this may introduce noise in the training phase of the Machine Learning model. To address this potential problem, we apply an initial filtering step in which only the gaits with the highest signal energy level are selected. Specifically, gait segments are sorted according to their AAV value, from highest to lowest. Notably, AAV is positively correlated with the signal's energy level. From this sorted list, only the first M gait segments are selected, while the others are discarded.

The retained M gait segments are used as inputs to the feature extraction process. Here, the most relevant aspects of the subject's gait pattern are captured by features computed in the time and time-frequency domains. Features and the feature selection procedure are described and discussed in Section 6.2.

Finally, a Machine Learning model is used to assess the frailty status. Let us define *gait instance* the vector of features derived from a gait segment. First, the NR or R class is determined for each one of the subject's gait instances. The subject is then categorized using a majority voting scheme, as indicated in Algorithm 6.

Algorithm 6: ML-based classification algorithm

Result: Subject's frailty status (R or NR)
Let $G = \{g_1, \ldots, g_M\}$ be the set of subject's gait instances;
foreach $g_i \in G$ **do**
  g_i is classified as R or NR;
end
Let $T = \{t_1, \ldots, t_M\}$, be the set of labels assigned by the classifier;
Let T_{NR} be the total number of gait segments labeled as NR;
Let k_{NR} be the non-robust parameter;
$threshold \leftarrow \lceil M \cdot k_{NR} \rceil$;
if $T_{NR} \geq threshold$ **then**
  Subject is classified as NR;
else
  Subject is classified as R;
end

The k_{NR} parameter was determined empirically. More details are presented in Section 6.2.

6.2 The method's validation criteria

In this section, we discuss the validation process of the proposed method, along with the characteristics of the experimental setup. In particular, we describe the subjects' inclusion procedure, the technology adopted during data collection, the

study of parameters and thresholds used in the process, and finally, the evaluation of frailty status assessment performance.

Participants

We enrolled adults aged 70+ admitted after accessing the Geriatrics outpatients clinic at Careggi academic hospital as patients or patients' caregivers. Subjects reporting physical dependence in at least one of Katz's basic activities of daily living (BADL) (Katz et al., 1963), as well as those with conditions causing overt abnormalities of gait (stroke, Parkinson's disease, severe hip or knee osteoarthritis) were excluded.

A total of 35 eligible subjects were included in the study, consisting of 14 females (78.86 ± 5.55 years old, height 1.61 ± 0.07 m, weight 68.50 ± 12.56 kg) and 21 males (80.00 ± 5.82 years old, height 1.70 ± 0.05 m, weight 77.81 ± 13.76 kg).

The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and was approved by the Local Ethics Committee*Comitato Etico Regionale per la Sperimentazione Clinica della Regione Toscana* (approval n. 14834_oss of May 7, 2019). A signed, written consent to participate in the study was obtained from all participants. Identifiable information was removed from the collected data to ensure participant anonymity.

Clinical assessment and experimental setup

The presence of frailty in participants was assessed, based on Fried's criteria, by measuring the following dimensions:

1. unintentional weight loss of 4.5 kg or more in the previous year;
2. low energy, identified through the CES-D (Center of Epidemiologic Studies Depression Scale) (Orme et al., 1986);
3. low physical activity defined thanks to the Physical Activity Questionnaire for the Elderly (PASE) (Schuit et al., 1997);
4. slowness, defined by the speed measured over a distance of 4.5 m and normalized for height and gender;
5. weakness, meaning reduced hand-grip strength in the dominant hand.

A subject was considered *robust*, *pre-frail* or *frail* if positive for three or more dimensions, one or two dimensions, or negative for all dimensions, respectively.

As a result, the sample contained 19 NR (non-robust, including 7 frail and 12 pre-frail) subjects and 16 R (robust) subjects.

After clinical evaluation, subjects were asked to wear a *Shimmer3* device on their wrists for a period of 24 hours, during which participants led their usual lives. Shimmer3 is a wearable device embedding a tri-axial accelerometer (STMicro LSM303DLHC) (Shimmer, 2018), which was used to collect acceleration samples at 102.4Hz. At the end of 24 hours, the subject returned to the clinic, and the sensor data were downloaded through the Shimmer Consensys software. We performed all further analysis on the signal on Python 3.9 notebooks, specifically created on *Jupyter Lab* (Granger and Pérez, 2021).

Subject Activity Level

As mentioned in Section 6.1, we introduced a measure that quantifies the level of activity of the subject during the day: the SAL value. In addition, as showed in Equation 6.1, we included in the SAL formula an α parameter that depends on the study context. In order to maximize the performance of our frailty status assessment in a home context, we tested all possible α values in the range $[0, 1]$, with an increment of 0.05. Specifically, we performed a dual investigation:

1. α-based statistical analysis of SAL;
2. α-based performance evaluation of the system.

In the first analysis, we aimed to identify the value of α that maximized the statistical significance of SAL in distinguishing between NR and R subjects. To this purpose, an *independent two-sample t-test* has been performed to compare the distributions of SAL values of R vs. NR subjects. The aim was to find the α which minimized the p-value in the latter comparison.

In the second analysis, we studied the performance of the entire frailty status assessment system as α varied, intending to identify the alpha that maximized R vs. NR classification scores. The results of this dual investigation are presented in Section 6.3.

Feature extraction

As previously mentioned, the gait segments identified in the segment labeling phase become the input to a feature extraction process. Here, a vector of features is computed from the gait segment's signal, including common statistical parameters used in signal processing (mean, median, standard deviation, minimum and maximum values, interquartile range (IQR), mean absolute deviation (MAD), root mean square (RMS), kurtosis, skewness and zero-crossing rate (ZCR)), calculated on acceleration components or Wavelet coefficients. Furthermore, two quantities previously used in gait analysis and fall detection studies are computed: the *cadence*, defined

Table 6.1: List of extracted features in the time domain for ecologic assessment of frailty using a wrist-worn device.

Feature	Components
Mean	x, y, z, m
Median	x, y, z, m
Standard deviation	x, y, z, m
Minimum value	x, y, z, m
Maximum value	x, y, z, m
Interquartile range	x, y, z, m
Kurtosis	x, y, z, m
Zero Crossing Rate	x, y, z, m
Mean Absolute Deviation	x, y, z, m
Root Mean Square	x, y, z, m
Average Absolute Variation	x, y, z, m
Cadence	-
Duration	-

as the ratio between the duration of the gait segment and the number of performed steps, and the average absolute acceleration variation (AAV), which is computed on consecutive acceleration samples (Cola et al., 2014, 2017).

The importance of walk-related wavelet features for frailty status assessment has already been shown in (Minici et al., 2021). In particular, we use the *Continuous Wavelet Transform* (CWT) on the acceleration magnitude signal to obtain a representation of the gait segment into the time-frequency domain. CWT is a signal processing technique to analyze a time series containing non-stationary power at different frequencies (Daubechies, 1990). More specifically, CWT allows the analysis of local variations of power by decomposing a time series into different frequency components.

The list of features included in the *gait instance*, for time and time-frequency domains, are shown in Tables 6.1 and 6.2, respectively.

Testing of the sensor-based frailty status assessment

A Leave-One-Subject-Out cross-validation (LOSO CV) procedure was used to test the frailty status assessment algorithm: at each iteration, the data (SAL and gait instances) of one subject were used as the testing set, while the data of other subjects were used as the training set. Besides, SAL values of the training set were used, at

Table 6.2: List of extracted features in the time-frequency domain for ecologic assessment of frailty using a wrist-worn device.

Feature	Components
Mean	CWT coefficients
Median	CWT coefficients
Standard deviation	CWT coefficients
Minimum value	CWT coefficients
Maximum value	CWT coefficients
Interquartile range	CWT coefficients
Kurtosis	CWT coefficients
Skewness	CWT coefficients

each iteration, to compute the AL_{NR} and AL_R thresholds introduced in Section 6.1. After finding the α that maximized the statistical significance of SAL, we compared the distributions of activity levels of R and NR subjects. From this comparison, we decided to assign to AL_{NR} the 25th percentile value of NR SALs, while we assigned to AL_R the 75th percentile of R SALs. More details on the above are provided in Section 6.3.

Three different Machine Learning models have been evaluated in the ML-based classification module: Random Forest, Gaussian naive Bayes, and Logistic Regression. All the classifiers have been implemented by means of the Python module Scikit-learn (Pedregosa et al., 2011).

In the ML-based classification module, a *feature selection* step was performed within the cross-validation procedure. Specifically, at each Leave-One-Subject-Out iteration, we applied a feature selection algorithm based on the One Way ANalysis Of VAriance (ANOVA) test to the training set. This algorithm computed the ANOVA F-statistic using values extracted from the R and NR subjects in the training set for each feature. Features were then sorted in descending order by F-statistic, and the best k features were selected. In this case, $k = 35$. The ML model was then trained using only the selected features of the training set and tested on the left-out subject.

From now on, let us consider NR subjects as positive and R subjects as negative classification results. We used *accuracy*, *sensitivity* and *specificity* scores to evaluate performance of the classification models. In particular, the accuracy score measures the ratio between the number of correctly classified NR and R subjects and the total number of subjects. Sensitivity, or true positive rate, measures the proportion of NR subjects that are correctly identified. On the other hand, the specificity score, also called the true negative rate, measures the proportion of R subjects correctly

identified. In addition, *Receiver Operating Characteristic* (ROC) *Area Under the Curve* (AUC) has been computed for each model.

6.3 Results and Discussion

The results obtained from the experiments described in Section 6.2 are now shown and discussed. First, we analyze the SAL and how we assigned the value to the α parameter. Next, we turn to the presentation of the results of the frailty status assessment algorithm.

Subject Activity Level

As already described in Section 6.2, we performed the following investigations on the α parameter of the SAL formula:

1. α-based statistical analysis of SAL;
2. α-based performance evaluation of the system.

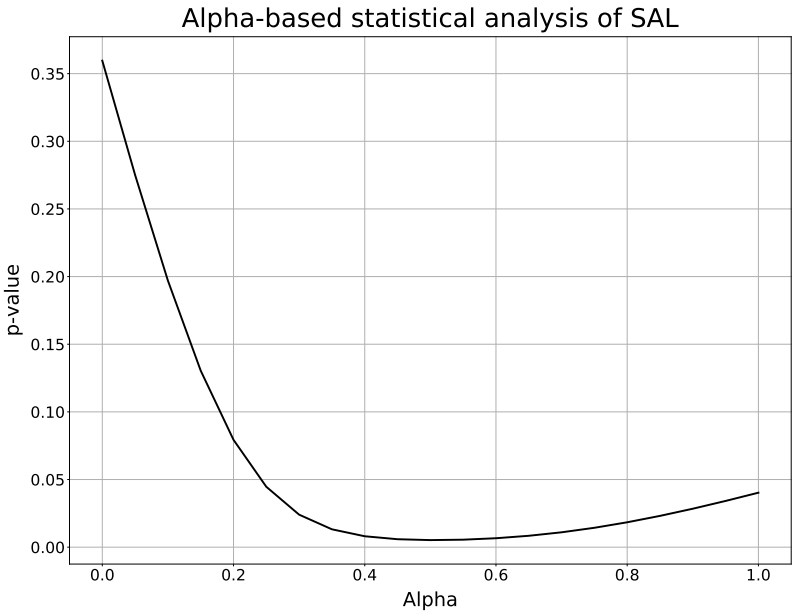


Figure 6.3: Results obtained in the α-based statistical analysis of the Subject Activity Level.

Figures 6.3 and 6.4 summarize the results of steps 1 and 2 of our α-based experiments, respectively. On the x-axis of the graph depicted in Figure 6.3, we find the

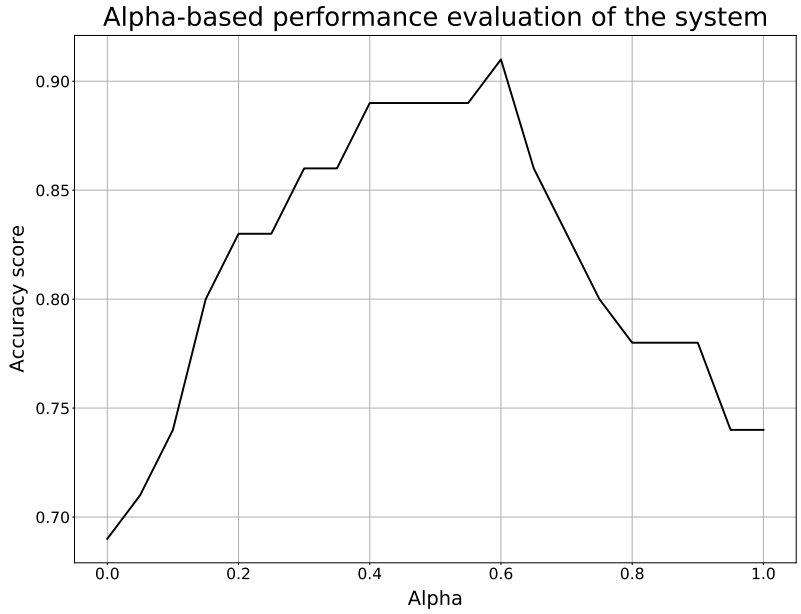


Figure 6.4: Results obtained in the α-based performance evaluation of the system.

alpha parameter of the SAL, while on the y-axis, we have the p-value calculated by the statistical t-test R SAL vs. NR SAL. It is important to remember how a lower p-value corresponds to a higher statistical significance of the SAL in distinguishing between R and NR subjects. In Figure 6.4, we again find alpha on the x-axis, but, in this case, the accuracy of the frailty status assessment algorithm is displayed on the y-axis. In other words, we can observe how the classification score of the system varies as alpha varies.

From the statistical analysis, we can notice how the highest statistical significance of SAL has been obtained for α values included in the range $[0.4, 0.6]$. Confirming this, the accuracy values of the frailty status assessment algorithm in the same interval are very positive, with a maximum of 0.91 for $\alpha = 0.6$.

Referring to the SAL formula we developed - $SAL = (1 - \alpha)ML + \alpha GL$ - choosing an α value in the above range means giving similar importance to walking and general movements performed during the day. On a quantitative level, it is indeed essential to consider all activities carried out by older adults, even those in which the person stands still without walking. Let us not forget that the dataset was collected during the COVID-19 pandemic. Thus, older subjects spent less time outside, and the actions performed were restricted to home environments. In other words, we may have recorded a limited number of walking intervals and other movements. That said, robust subjects may still have had a higher activity level than non-robust

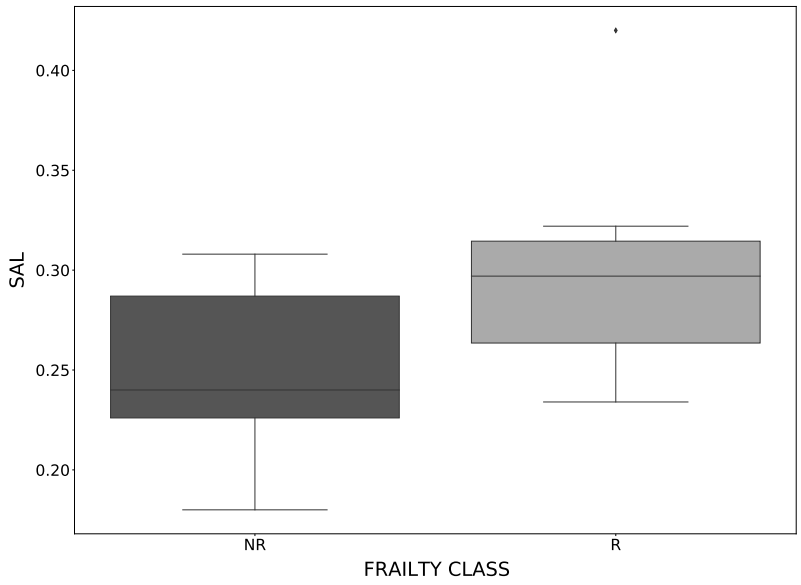


Figure 6.5: Box plot representation of the Subject Activity Level computed from the R and NR subjects.

subjects, which in the case of our study, are captured by an almost equally balanced SAL.

In light of what we obtained from our investigation, we assigned $\alpha = 0.6$, and the SAL formula became as follows:

$$SAL = 0.4ML + 0.6GL,$$

Before presenting the results of the frailty status assessment algorithm, let us end the discussion regarding the SAL by examining how we selected the AL_{NR} and AL_R thresholds. Figure 6.5 shows the box plots generated by SAL values of NR and R subjects. From a simple observation of the plot, we note that all R subjects have a SAL greater than the first quartile of NR SALs. Likewise, all NR subjects have a SAL lower than the third quartile of R SALs. For this reason, we decided to assign to AL_{NR} and AL_R these two values, computed from the R and NR subjects in the training set, at each LOSO CV iteration. This way, the evaluation procedure and the results obtained were kept valid.

Frailty status assessment

We collected data from 35 older adults over a 24-hour time interval, during which the subjects lived their everyday lives. The data were then used throughout the

Table 6.3: Average results of the frailty status assessment algorithm.

Classifier	Acc.	Sens.	Spec.	AUC
Gaussian Naïve Bayes	0.91	0.94	0.88	0.91
Random Forest	0.77	0.68	0.88	0.78
Logistic Regression	0.74	0.74	0.75	0.74

process described in the 6.1 section. The frailty status assessment algorithm was adapted to the study context by tuning the parameters and thresholds defined in the 6.2 section. Finally, we tested the entire system through a LOSO CV procedure. Table 6.3 summarizes the results obtained for all three classifiers tested.

The Gaussian Naïve Bayes model performed substantially better than the other two models, classifying subjects with an accuracy score of 91% (32 correctly classified participants out of 35). In addition, the algorithm recognized 18 out of 19 NR subjects (94% sensitivity score) and 14 out of 16 R subjects (88% specificity score). The Random Forest model achieved higher accuracy than Logistic Regression, though it achieved a lower sensitivity score, which is a crucial indicator in healthcare-related applications (low sensitivity leads to false negatives, i.e., frail patients are missed by the system). The latter model had the worst classification performance, although the scores were very balanced.

These results bring several positive aspects to consider, in line with the aims discussed in the Introduction. First, we verified what was already shown in (Minici et al., 2021). However, unlike the previous study, we collected sensor-derived signals in an unsupervised setting, and still the results were remarkable. This achievement further supports the hypothesis that sensor-based gait biomarkers, in combination with a Machine Learning-based algorithm, may enable a practical approach to continuous frailty status monitoring.

Second, this chapter confirmed the wrist as an excellent body position for collecting frailty-related accelerometer signals. Other researchers have explored the use of sensors to monitor frail patients. However, the procedures used were less practical, as sensors were attached to a t-shirt or to the user's lower back employing special belts. Compared to these previous experiences, a wrist-worn device is more convenient and unobtrusive. This might improve the compliance of older adults asked to wear a device for an extended time interval. Again, our results represent a significant step towards the feasibility of automated frailty status assessment.

Finally, in Chapter 1, we mentioned the correlation between frailty and physical activity and discussed how this correlation is often studied through self-reported questionnaires. In our study, we defined the objective measure SAL and used this index jointly with the *ML-based classification* module. Our goal was to combine a

quantitative measure that collects information about the subject's entire day with a qualitative measure that finds specific gait-related details in short time intervals. The achieved results support the validity of this approach and the correlation between a sedentary life and frailty condition. Again, these findings could represent a further step toward improving monitoring systems aimed at frailty status assessment.

Chapter 7

Conclusions

In this dissertation, we explored the use of wearable sensors combined with mathematical models and artificial intelligence tools for frailty identification.

First, we investigated the use of wearable sensors to analyze gait and identify pre-frail and frail subjects. To this purpose, gait segments from 34 older adults were detected by using two wearable devices, placed at the subject's wrist and lower back. These traces were then independently analyzed, in time and time-frequency domains. Five different machine learning models were tested for the classification task: Random Forest, Gaussian naive Bayes, Logistic Regression, Multilayer Perceptron and Support Vector Machine. The best performance was achieved by Gaussian naive Bayes (91.3% sensitivity and 81.8% specificity) using a combination of time and time-frequency domains features extracted from a wrist sensor-derived signal. Interestingly, the wrist-worn sensor achieved substantially better classification accuracy than the lower-back-worn sensor, indicating arm motion as a distinctive element between robust and frail subjects. Continuous Wavelet Analysis was used to include information about the distribution of power in the time-frequency domain, that allowed us to capture gait regularity and achieve better classification performance. The results demonstrated that unobtrusive wearable devices may enable an effective approach in continuous monitoring of human walking, and represent a significant step towards the feasibility of automated frailty status assessment based on machine-learning. Also, from the application point of view, a wrist-worn based implementation of the proposed method may foster user adoption of wearable devices for early detection of the frailty syndrome, as it may be embedded into a smartwatch.

Secondly, we analyzed the output of the Continuous Wavelet Transform applied to the acceleration during gait, in order to identify the most useful CWT-based frequency bands (FB) for frailty identification. To this purpose, the gait traces previously collected from the 34 older adults were used. A time-frequency domain representation of these traces was then obtained by means of CWT. Features were computed on seven different frequency bands of the CWT. Statistical differences be-

tween the features of the three groups (i.e., robust, pre-frail, frail) were analyzed using ANOVA and t-tests. In addition, the feature sets were used to train and test seven different random forest classifiers. The best performance was achieved by the classifier trained with the gait instances of $[1.5, 2.5]Hz$ FB, which also represents the most statistically significant FB for distinguishing subjects according to their frailty status. These results suggested that features extracted in $[1.5, 2.5]Hz$ band may improve the performance of frailty status assessment models, as they contain statistically significant frailty-related information.

Next, we used a Deep Convolutional Neural Network to further analyze the output of Continuous Wavelet Transform, since the scalograms extracted from subjects belonging to the same frailty class seem to present similar patterns. In particular, CWT was applied to the signal traces extracted during a walk activity of 34 older adults, producing scalograms as output. A Convolutional Neural Network was trained in two different stages, by applying a Transfer Learning strategy. We trained a base model for Human Activity Recognition purposes based on the scalograms obtained from a public dataset. A second-stage training was then performed on the last layers of the CNN, this time for frailty status classification using the scalograms extracted from the 34 subjects. The results demonstrated that Deep Learning-based systems might represent an effective approach to the frailty classification task, although they did not improve the scores previously achieved using shallow learning models.

Finally, we explored the use of in-home collected wearable-derived signals for frailty status assessment. To this aim, 24-hour acceleration traces were extracted from 35 subjects aged 70+ using a wrist-worn device. We processed these signals and trained a frailty status assessment algorithm based on subjects' activity levels and Machine Learning classification models. We tested three Machine Learning classifiers: Random Forest, Gaussian naive Bayes, and Logistic Regression. The best performance was achieved with the Gaussian naive Bayes used as the classifier in the ML-based module (94% sensitivity and 88% specificity). These findings gain relevance if one considers that participants wore the device at the wrist. A wrist-worn device is more practical and can increase the compliance of older adults to use wearable devices for continuous monitoring of their clinical condition. Indeed, the proposed algorithm may be integrated into a commercial smartwatch. Finally, we demonstrated how the subjects' activity levels could characterize robust and non-robust subjects. We combined an objectively measured quantity with Machine Learning-derived outcomes to classify subjects according to their frailty status. This helped improve the achieved scores, highlighting the correlation between activity level and frailty.

7.1 Limitations and Future Work

The limited number of subjects involved in the experiments represents a limitation of this thesis. A second limitation is that we contrasted robust with non-robust subjects, whereas we could not reproduce the original three-level classification of frailty status. In future work, we plan to expand the dataset in order to test the generalizability of our Machine Learning models in a larger and more diverse sample. This would also consent to distinguish between pre-frail and frail individuals.

Even with these limitations, our results demonstrate that sensor-based gait biomarkers and a Machine Learning-based algorithm represent a suitable and valid approach for automated, ecologic assessment of frailty status monitoring. From a future perspective, it might be envisioned that this approach would eventually consent also long-term monitoring of older persons in their homes to detect early deviation from robustness towards frailty.

Appendix A

Ph.D. Internship at Novartis Institutes for BioMedical Research

As part of the Smart Computing Ph.D. program, I was a Ph.D. candidate at the Novartis Institutes for BioMedical Research in Basel (Switzerland). I worked as a data scientist in Translational Medicine, Biomarker Development, Quantitative Sciences and Innovation. The work focused on mobile medical devices, very large multivariate clinical datasets, and real-world data. In particular, I was involved in two different projects. The first project had the goal to select the key gait parameters, extracted by a network of accelerometers, in the motor-cognitive interference in dual task paradigm. This project directly informed the company research conducted on diseases affected by cognitive decline. The second project aimed to explore the importance of physical activity metrics obtained from several wearable devices as predictor of changes and treatment efficacy in different clinical conditions.

Appendix B

Publications

Journal papers

1. **D. Minici**, G. Cola, A. Giordano, S. Antoci, E. Girardi, M. Di Bari, M. Avvenuti, "Towards automated assessment of frailty status using a wrist-worn device", *IEEE Journal of Biomedical and Health Informatics*, pages: 1013–1022, 2021.
   Candidate's contributions: conceptualization, sensor preparation, data extraction and analysis, investigation, methodology, software, validation, visualization, writing - original draft, writing - review & editing.

2. M.G.C.A. Cimino, **D. Minici**, M. Monaco, S. Petrocchi, G. Vaglini, "A hyper-heuristic methodology for coordinating swarms of robots in target search", *Computers and Electrical Engineering*, pages: 107420, 2021.
   Candidate's contributions: investigation, simulation, methodology, software, validation, writing - review & editing.

Peer reviewed conference papers

1. **D. Minici**, G. Cola, A. Giordano, S. Antoci, E. Girardi, M. Di Bari, M. Avvenuti, "Wavelet-based analysis of gait for automated frailty assessment with a wrist-worn device", *2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages: 1–4, 2021.
   Candidate's contributions: conceptualization, sensor preparation, data extraction and analysis, investigation, methodology, software, validation, visualization, writing - original draft, writing - review & editing.

Papers under review

1. **D. Minici**, G. Cola, G. Perfetti, S. Espinoza Tofalos, M. Di Bari, M. Avvenuti, "Automated, ecologic assessment of frailty using a wrist-worn device", *IEEE Journal of Biomedical and Health Informatics*, 2022.
   Candidate's contributions: conceptualization, sensor preparation, data extraction and analysis, investigation, methodology, software, validation, visualization, writing - original draft, writing - review & editing.

Other

1. **D. Minici**, G. Cola, A. Giordano, S. Antoci, E. Girardi, M. Di Bari, M. Avvenuti, "Sensor-based assessment of gait to evaluate frailty status in older adults: preliminary data from the WeSPA study.", *6th Italian Conference on ICT for Smart Cities and Communities* (*I-CiTies 2020*).
   Candidate's contributions: conceptualization, sensor preparation, data extraction and analysis, investigation, methodology, software, validation, visualization, writing - original draft, writing - review & editing, conference presentation.

2. **D. Minici**, A. Giordano, G. Cola, S. Antoci, E. Girardi, M. Di Bari, M. Avvenuti, "Wearable-Sensors based Personalized Assessment (WeSPA) of frailty: preliminary findings", *2020 European Geriatric Medicine Society EuGMS*.
   Candidate's contributions: conceptualization, sensor preparation, data extraction and analysis, investigation, methodology, software, validation, visualization.

3. **D. Minici**, G. Perfetti, S. Espinoza Tofalos, G. Cola, I. Ambrosino, M. Avvenuti, M. Di Bari, "Sensori indossabili per lo screening della fragilità nell'anziano: il progetto WeSPA" *2022 Congresso Nazionale della Società Italiana di Gerontologia e Geriatria SIGG*.
   Candidate's contributions: conceptualization, sensor preparation, data extraction and analysis, investigation, methodology, software, validation, visualization.

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Abbate, S., Avvenuti, M., and Light, J. (2012). Mims: A minimally invasive monitoring sensor platform. *IEEE Sensors Journal*, 12(3):677–684.

Abizanda, P., Romero, L., Sanchez-Jurado, P., Atienzar-Nunez, P., Esquinas-Requena, J., and García-Nogueras, I. (2012). Association between functional assessment instruments and frailty in older adults: The fradea study. *The Journal of frailty & aging*, 1(4):162–168.

Albahri, O. S., Zaidan, A., Zaidan, B., Hashim, M., Albahri, A. S., and Alsalem, M. (2018). Real-time remote health-monitoring systems in a medical centre: A review of the provision of healthcare services-based body sensor information, open challenges and methodological aspects. *Journal of medical systems*, 42(9):1–47.

Antonsson, E. K. and Mann, R. W. (1985). The frequency content of gait. *Journal of biomechanics*, 18(1):39–47.

Avvenuti, M., Carbonaro, N., Cimino, M. G., Cola, G., Tognetti, A., and Vaglini, G. (2018). Smart shoe-assisted evaluation of using a single trunk/pocket-worn accelerometer to detect gait phases. *Sensors*, 18(11):3811.

Bergman, H., Béland, F., Feightner, J., et al. (2003). The canadian initiative on frailty and aging. *Aging Clin Exp Res*, 15(3 Suppl):1–2.

Bieńkiewicz, M. M., Brandi, M.-L., Goldenberg, G., Hughes, C. M., and Hermsdörfer, J. (2014). The tool in the brain: apraxia in adl. behavioral and neurological correlates of apraxia in daily living. *Frontiers in psychology*, 5:353.

Bock, J.-O., König, H.-H., Brenner, H., Haefeli, W. E., Quinzler, R., Matschinger, H., Saum, K.-U., Schöttker, B., and Heider, D. (2016). Associations of frailty with health care costs–results of the esther cohort study. *BMC health services research*, 16(1):1–11.

Bruijn, S. M., Meijer, O. G., Beek, P. J., and van Dieën, J. H. (2010). The effects of arm swing on human gait stability. *Journal of experimental biology*, 213(23):3945–3952.

Chollet, F. et al. (2015). Keras. `https://keras.io`.

Cola, G., Avvenuti, M., Musso, F., and Vecchio, A. (2017). Personalized gait detection using a wrist-worn accelerometer. In *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 173–177.

Cola, G., Avvenuti, M., Piazza, P., and Vecchio, A. (2016). Fall detection using a head-worn barometer. In *International Conference on Wireless Mobile Communication and Healthcare*, pages 217–224. Springer.

Cola, G., Vecchio, A., and Avvenuti, M. (2014). Improving the performance of fall detection systems through walk recognition. *Journal of Ambient Intelligence and Humanized Computing*, 5(6):843–855.

Collard, R. M., Boter, H., Schoevers, R. A., and Oude Voshaar, R. C. (2012). Prevalence of frailty in community-dwelling older persons: a systematic review. *Journal of the American Geriatrics Society*, 60(8):1487–1492.

Dasenbrock, L., Heinks, A., Schwenk, M., and Bauer, J. (2016). Technology-based measurements for screening, monitoring and preventing frailty. *Zeitschrift für Gerontologie und Geriatrie*, 49(7):581–595.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, 36(5):961–1005.

Dent, E., Martin, F. C., Bergman, H., Woo, J., Romero-Ortuno, R., and Walston, J. D. (2019). Management of frailty: opportunities, challenges, and future directions. *The Lancet*, 394(10206):1376–1386.

Díaz, S., Stephenson, J. B., and Labrador, M. A. (2020). Use of wearable sensor technology in gait, balance, and range of motion analysis. *Applied Sciences*, 10(1):234.

Dipietro, L., Campbell, W. W., Buchner, D. M., Erickson, K. I., Powell, K. E., Bloodgood, B., Hughes, T., Day, K. R., Piercy, K. L., Vaux-Bjerke, A., et al. (2019). Physical activity, injurious falls, and physical function in aging: an umbrella review. *Medicine and science in sports and exercise*, 51(6):1303.

Edemekong, P. F., Bomgaars, D. L., Sukumaran, S., and Levy, S. B. (2021). Activities of daily living. In *StatPearls* [*Internet*]. StatPearls Publishing.

Fairhall, N., Sherrington, C., Kurrle, S. E., Lord, S. R., Lockwood, K., Howard, K., Hayes, A., Monaghan, N., Langron, C., Aggar, C., et al. (2015). Economic evaluation of a multifactorial, interdisciplinary intervention versus usual care to reduce frailty in frail older people. *Journal of the American Medical Directors Association*, 16(1):41–48.

Fried, L. P., Herdman, S. J., Kuhn, K. E., Rubin, G., and Turano, K. (1991). Preclinical disability: hypotheses about the bottom of the iceberg. *Journal of Aging and Health*, 3(2):285–300.

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G., et al. (2001). Frailty in older adults: evidence for a phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3):M146–M157.

Galán-Mercant, A. and Cuesta-Vargas, A. I. (2013). Differences in trunk accelerometry between frail and nonfrail elderly persons in sit-to-stand and stand-to-sit transitions based on a mobile inertial sensor. *JMIR mhealth and uhealth*, 1(2):e21.

García-Esquinas, E., Andrade, E., Martínez-Gómez, D., Caballero, F. F., López-García, E., and Rodríguez-Artalejo, F. (2017). Television viewing time as a risk factor for frailty and functional limitations in older adults: results from 2 european prospective cohorts. *International Journal of Behavioral Nutrition and Physical Activity*, 14(1):1–18.

García-Villamil, G., Neira-Álvarez, M., Huertas-Hoyas, E., Ramón-Jiménez, A., and Rodríguez-Sánchez, C. (2021). A pilot study to validate a wearable inertial sensor for gait assessment in older adults with falls. *Sensors*, 21(13):4334.

Garcìa-Nogueras, I., Aranda-Reneo, I., Peña-Longobardo, L., Oliva-Moreno, J., and Abizanda, P. (2017). Use of health resources and healthcare costs associated with frailty: the fradea study. *The journal of nutrition, health & aging*, 21(2):207–214.

Gilpin, R. (2018). The challenge of global capitalism. In *The Challenge of Global Capitalism*. Princeton University Press.

Gobbens, R. J., Luijkx, K. G., Wijnen-Sponselee, M. T., and Schols, J. M. (2010). Toward a conceptual definition of frail community dwelling older people. *Nursing outlook*, 58(2):76–86.

Granger, B. and Pérez, F. (2021). Jupyter: Thinking and storytelling with code and data. *Authorea Preprints*.

Greene, B. R., Doheny, E. P., O'Halloran, A., and Anne Kenny, R. (2014). Frailty status can be accurately assessed using inertial sensors and the tug test. *Age and ageing*, 43(3):406–411.

Guralnik, J. M., Ferrucci, L., Simonsick, E. M., Salive, M. E., and Wallace, R. B. (1995). Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *New England Journal of Medicine*, 332(9):556–562.

Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., Scherr, P. A., and Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *Journal of gerontology*, 49(2):M85–M94.

Hantke, N. C. and Gould, C. (2020). Examining older adult cognitive status in the time of covid-19. *Journal of the American Geriatrics Society*.

Huisingh-Scheetz, M., Kocherginsky, M., Dugas, L., Payne, C., Dale, W., Conroy, D. E., and Waite, L. (2016). Wrist accelerometry in the health, functional, and social assessment of older adults. *Journal of the American Geriatrics Society*, 64(4):889.

Huisingh-Scheetz, M., Wroblewski, K., Kocherginsky, M., Huang, E., Dale, W., Waite, L., and Schumm, L. P. (2018). The relationship between physical activity and frailty among us older adults based on hourly accelerometry data. *The Journals of Gerontology: Series A*, 73(5):622–629.

Katz, S. (1983). Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *Journal of the American Geriatrics Society*.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., and Jaffe, M. W. (1963). Studies of illness in the aged: the index of adl: a standardized measure of biological and psychosocial function. *Jama*, 185(12):914–919.

Kehler, D. S., Hay, J. L., Stammers, A. N., Hamm, N. C., Kimber, D. E., Schultz, A. S., Szwajcer, A., Arora, R. C., Tangri, N., and Duhamel, T. A. (2018). A systematic review of the association between sedentary behaviors with frailty. *Experimental gerontology*, 114:1–12.

Kehler, D. S. and Theou, O. (2019). The impact of physical activity and sedentary behaviors on frailty levels. *Mechanisms of ageing and development*, 180:29–41.

Khandelwal, S. and Wickström, N. (2016). Gait event detection in real-world environment for long-term applications: Incorporating domain knowledge into time-frequency analysis. *IEEE transactions on neural systems and rehabilitation engineering*, 24(12):1363–1372.

Kosse, N. M., Vuillerme, N., Hortobágyi, T., and Lamoth, C. J. (2016). Multiple gait parameters derived from ipod accelerometry predict age-related gait changes. *Gait & posture*, 46:112–117.

Mañas, A., del Pozo-Cruz, B., García-García, F. J., Guadalupe-Grau, A., and Ara, I. (2017). Role of objectively measured sedentary behaviour in physical performance, frailty and mortality among older adults: A short systematic review. *European journal of sport science*, 17(7):940–953.

Manfredi, G., Midão, L., Paúl, C., Cena, C., Duarte, M., and Costa, E. (2019). Prevalence of frailty status among the european elderly population: Findings from the survey of health, aging and retirement in europe. *Geriatrics & gerontology international*, 19(8):723–729.

Martínez-Ramírez, A., Lecumberri, P., Gómez, M., Rodriguez-Mañas, L., García, F., and Izquierdo, M. (2011). Frailty assessment based on wavelet analysis during quiet standing balance test. *Journal of biomechanics*, 44(12):2213–2220.

Martínez-Ramírez, A., Martinikorena, I., Gómez, M., Lecumberri, P., Millor, N., Rodríguez-Mañas, L., García, F. J. G., and Izquierdo, M. (2015). Frailty assessment based on trunk kinematic parameters during walking. *Journal of Neuroengineering and rehabilitation*, 12(1):1–10.

Millor, N., Lecumberri, P., Gómez, M., Martínez-Ramírez, A., Rodríguez-Mañas, L., García-García, F. J., and Izquierdo, M. (2013). Automatic evaluation of the 30-s chair stand test using inertial/magnetic-based technology in an older prefrail population. *IEEE journal of biomedical and health informatics*, 17(4):820–827.

Minici, D., Cola, G., Giordano, A., Antoci, S., Girardi, E., Di Bari, M., and Avvenuti, M. (2021). Towards automated assessment of frailty status using a wrist-worn device. *IEEE Journal of Biomedical and Health Informatics*, 26(3):1013–1022.

Mirelman, A., Bernad-Elazari, H., Nobel, T., Thaler, A., Peruzzi, A., Plotnik, M., Giladi, N., and Hausdorff, J. M. (2015). Effects of aging on arm swing during gait: the role of gait speed and dual tasking. *PLoS One*, 10(8):e0136043.

Moe-Nilssen, R. and Helbostad, J. L. (2004a). Estimation of gait cycle characteristics by trunk accelerometry. *Journal of Biomechanics*, 37(1):121 – 126.

Moe-Nilssen, R. and Helbostad, J. L. (2004b). Estimation of gait cycle characteristics by trunk accelerometry. *Journal of biomechanics*, 37(1):121–126.

Mohler, M. J., Fain, M. J., Wertheimer, A. M., Najafi, B., and Nikolich-Žugich, J. (2014). The frailty syndrome: clinical measurements and basic underpinnings in humans and animals. *Experimental gerontology*, 54:6–13.

Montero-Odasso, M., Muir, S. W., Hall, M., Doherty, T. J., Kloseck, M., Beauchet, O., and Speechley, M. (2011). Gait variability is associated with frailty in community-dwelling older adults. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 66(5):568–576.

Orme, J. G., Reis, J., and Herz, E. J. (1986). Factorial and discriminant validity of the center for epidemiological studies depression (ces-d) scale. *Journal of clinical psychology*, 42(1):28–33.

Pashmdarfard, M. and Azad, A. (2020). Assessment tools to evaluate activities of daily living (adl) and instrumental activities of daily living (iadl) in older adults: A systematic review. *Medical journal of the Islamic Republic of Iran*, 34:33.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830.

Pradeep Kumar, D., Toosizadeh, N., Mohler, J., Ehsani, H., Mannier, C., and Laksari, K. (2020). Sensor-based characterization of daily walking: a new paradigm in pre-frailty/frailty assessment. *BMC geriatrics*, 20:1–11.

Pritchard, J., Kennedy, C., Karampatos, S., Ioannidis, G., Misiaszek, B., Marr, S., Patterson, C., Woo, T., and Papaioannou, A. (2017). Measuring frailty in clinical practice: a comparison of physical frailty assessment methods in a geriatric outpatient clinic. *BMC geriatrics*, 17(1):264.

Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., and Yang, G.-Z. (2016). Deep learning for health informatics. *IEEE journal of biomedical and health informatics*, 21(1):4–21.

Rezvanian, S. and Lockhart, T. E. (2016). Towards real-time detection of freezing of gait using wavelet transform on wireless accelerometer data. *Sensors*, 16(4):475.

Ritt, M., Schülein, S., Lubrich, H., Bollheimer, L., Sieber, C., and Gassmann, K.-G. (2017). High-technology based gait assessment in frail people: associations between spatio-temporal and three-dimensional gait characteristics with frailty status across four different frailty measures. *The journal of nutrition, health & aging*, 21(3):346–353.

Salinas-Rodrìguez, A., Manrique-Espinoza, B., Heredia-Pi, I., Rivera-Almaraz, A., and Avila-Funes, J. A. (2019). Healthcare costs of frailty: implications for long-term care. *Journal of the American Medical Directors Association*, 20(1):102–103.

Sanders, T. H., Devergnas, A., Wichmann, T., and Clements, M. A. (2013). Remote smartphone monitoring for management of parkinson's disease. In *Proceedings of*

the 6th International Conference on PErvasive Technologies Related to Assistive Environments, pages 1–5.

Schuit, A. J., Schouten, E. G., Westerterp, K. R., and Saris, W. H. (1997). Validity of the physical activity scale for the elderly (pase): according to energy expenditure assessed by the doubly labeled water method. *Journal of clinical epidemiology*, 50(5):541–546.

Schwenk, M., Howe, C., Saleh, A., Mohler, J., Grewal, G., Armstrong, D., and Najafi, B. (2014). Frailty and technology: a systematic review of gait analysis in those with frailty. *Gerontology*, 60(1):79–89.

Schwenk, M., Mohler, J., Wendel, C., Fain, M., Taylor-Piliae, R., Najafi, B., et al. (2015). Wearable sensor-based in-home assessment of gait, balance, and physical activity for discrimination of frailty status: baseline results of the arizona frailty cohort study. *Gerontology*, 61(3):258–267.

Shimmer (2018). http://www.shimmersensing.com.

Simpson, K. N., Seamon, B. A., Hand, B. N., Roldan, C. O., Taber, D. J., Moran, W. P., and Simpson, A. N. (2018). Effect of frailty on resource use and cost for medicare patients. *Journal of comparative effectiveness research*, 7(8):817–825.

Sirven, N. and Rapp, T. (2017). The cost of frailty in france. *The European Journal of health economics*, 18(2):243–253.

Tao, W., Liu, T., Zheng, R., and Feng, H. (2012). Gait analysis using wearable sensors. *Sensors*, 12(2):2255–2283.

Thiede, R., Toosizadeh, N., Mills, J. L., Zaky, M., Mohler, J., and Najafi, B. (2016). Gait and balance assessments as early indicators of frailty in patients with known peripheral artery disease. *Clinical biomechanics*, 32:1–7.

Tolley, A. P., Ramsey, K. A., Rojer, A. G., Reijnierse, E. M., and Maier, A. B. (2021). Objectively measured physical activity is associated with frailty in community-dwelling older adults: A systematic review. *Journal of clinical epidemiology*, 137:218–230.

Torrence, C. and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, 79(1):61–78.

van Kan, G. A., Rolland, Y. M., Morley, J. E., and Vellas, B. (2008). Frailty: toward a clinical definition. *Journal of the American Medical Directors Association*, 9(2):71–72.

Visser, M., Deeg, D. J., and Lips, P. (2003). Low vitamin d and high parathyroid hormone levels as determinants of loss of muscle strength and muscle mass (sarcopenia): the longitudinal aging study amsterdam. *The Journal of Clinical Endocrinology & Metabolism*, 88(12):5766–5772.

Weiss, G. M., Yoneda, K., and Hayajneh, T. (2019). Smartphone and smartwatch-based biometrics using activities of daily living. *IEEE Access*, 7:133190–133202.

Weiss, K., Khoshgoftaar, T. M., and Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3(1):1–40.

Xue, Q.-L. (2011). The frailty syndrome: definition and natural history. *Clinics in geriatric medicine*, 27(1):1–15.

Zacharaki, E. I., Deltouzos, K., Kalogiannis, S., Kalamaras, I., Bianconi, L., Degano, C., Orselli, R., Montesa, J., Moustakas, K., Votis, K., et al. (2020). Frailsafe: An ict platform for unobtrusive sensing of multi-domain frailty for personalized interventions. *IEEE Journal of Biomedical and Health Informatics*.

Zhang, W., Regterschot, G. R. H., Geraedts, H., Baldus, H., and Zijlstra, W. (2015). Chair rise peak power in daily life measured with a pendant sensor associates with mobility, limitation in activities, and frailty in old people. *IEEE journal of biomedical and health informatics*, 21(1):211–217.

Zhong, R., Rau, P.-L. P., and Yan, X. (2018). Application of smart bracelet to monitor frailty-related gait parameters of older chinese adults: A preliminary study. *Geriatrics & gerontology international*, 18(9):1366–1371.

Ziller, C., Braun, T., and Thiel, C. (2020). Frailty phenotype prevalence in community-dwelling older adults according to physical activity assessment method. *Clinical Interventions in Aging*, 15:343.

Zivanovic, M., Millor, N., and Gómez, M. (2018). Modeling of noisy acceleration signals from quasi-periodic movements for drift-free position estimation. *IEEE journal of biomedical and health informatics*, 23(4):1558–1565.