Pervasive and Mobile Computing 95 (2023) 101833

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

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

Domenico Minici [a,*], Guglielmo Cola [b], Giulia Perfetti [c,d], Sofia Espinoza Tofalos [c,d], Mauro Di Bari [c,d], Marco Avvenuti [a]

[a] Department of Information Engineering, University of Pisa, Pisa, Italy
[b] Institute of Informatics and Telematics, IIT-CNR, Pisa, Italy
[c] Research Unit of Medicine of Aging, Department of Clinical and Experimental Medicine, University of Florence, Firenze, Italy
[d] Unit of Geriatrics - Geriatrics Intensive Care Unit, Department of Medicine and Geriatrics, Careggi Hospital, Firenze, Italy

A R T I C L E I N F O

Article history:
Received 10 January 2023
Received in revised form 27 June 2023
Accepted 22 August 2023
Available online 23 August 2023

Keywords:
Frailty
Gait analysis
Machine learning
Wearable sensor
Wavelet transform

A B S T R A C T

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. This study aims to explore in-home collected wearable-derived signals for frailty status assessment. A sample of 35 subjects aged 70+, autonomous in basic activities of daily living and cognitively intact, was collected. After being clinically assessed for frailty according to Fried's phenotype, participants wore a wrist device equipped with inertial motion sensors for 24 h, during which they led their usual life in their homes. Signal-derived traces were split into 10-s segments and labeled classified as gaits, other motor activities, or rests. Gait and other motor activity segments were used to calculate the Subject Activity Level (SAL), 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. In particular, subjects were classified as robust or non-robust, a category that includes both Fried's frail and pre-frail phenotypes. For some users, activity levels alone enabled accurate frailty assessment, whereas, for others, a Gaussian Naive Bayes classifier based on the gait-derived features was required to assess frailty status. Overall, the proposed method showed extremely promising results, allowing discrimination of robust and non-robust subjects with an overall 91% accuracy, stemming from 95% sensitivity and 88% specificity. This study demonstrates the potential of unobtrusive, wearable devices in objectively assessing frailty through unsupervised monitoring in real-world settings.

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 [1]. 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

* Corresponding author.
E-mail address: domenico.minici@phd.unipi.it (D. Minici).

https://doi.org/10.1016/j.pmcj.2023.101833

care have already gained considerable attention, particularly for patients in isolated communities and remote regions [2,3]. 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 [1,4].

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 [5]. Among the conditions that undermine the quality of life of older persons, frailty represents one of the most severe global public health challenges [6]. Frailty is a condition of the reduced homeostatic reserve, which makes some older individuals more susceptible to endogenous and external stressors and increases their risk of adverse health outcomes, [7], including the development of disability, i.e., inability to perform Activities of Daily Living (ADL) [8,9]. The concept of frailty has been operationalized with several diagnostic tools, among which the phenotype model, developed by Fried et al. [10], is one of the most accepted. The following five items are considered for its application: unintentional weight loss, self-reported exhaustion, low energy expenditure, slow gait speed, and weak grip strength. Scoring positive in three or more, one or two, or none of these items classifies older subjects as frail, pre-frail, or robust, respectively. It can be easily argued that motor activities, and gait, in particular, exert a central role in the ascertainment of the frailty phenotype.

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) [8,9]. The relationship between frailty and ADL has been widely discussed in literature [11–15]. Dipietro et al. [13] suggest that regular physical activity effectively helps older adults delay the loss of mobility while reducing the risk of fall-related injuries. Authors in [15] 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 [16], whose results emphasize the importance of physical activities and their assessment methods.

In spite of its apparent simplicity, the application of Fried's phenotype requires personnel with adequate training and expertise, as well as space and time resources. All of these may lack in the busy routine of clinical practice and, as discussed above, were almost completely unavailable during the COVID-19 outbreak. Moreover, it might be hypothesized that a classification of frailty status based on an extended session of data collection, during which subjects perform their daily tasks in their homes, would be more accurate than the short assessment commonly employed to apply Fried's tool. Due to these factors, several researchers have focused on using wearable, sensor-based technologies to collect motor condition measures and produce an ecological and valid assessment of frailty in older adults [16–21]. In [22], 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. More recently, authors in [23] 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 [24] 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 mounting body of evidence has underscored the value of wrist-worn wearable technologies for assessing frailty. Huisingh-Scheetz et al. [25] utilized wrist-worn devices to gather hourly accelerometry data, demonstrating that activity measures could significantly distinguish between different frailty classes among older adults. Concurrently, Toosizadeh et al. [26] proposed an innovative method of assessing upper-extremity frailty through wearable sensors attached to the arm, uncovering a significant correlation between their frailty test and the Trauma-Specific Frailty Index. Similarly, Mulasso et al. [27] confirmed the validity of wrist-based accelerometry data for assessing frailty and found a strong correlation with conventional physical frailty assessments. Furthermore, Lee et al. [28] demonstrated the potential utility of wrist-wearable devices in evaluating frailty among bedridden hospital inpatients, underscoring the role of smartwatch-based assessments in enhancing patient care within clinical environments. However, the researchers noted the necessity for further exploring wearable technology applications in real-world settings.

The integration of machine learning and handcrafted features in healthcare, particularly for frailty assessment, is a rapidly evolving field. Ravi et al. [29] showcased machine learning techniques adapted for resource-constrained platforms such as wearable devices. Tarekegn et al. [30] emphasized the effectiveness of machine learning in predicting frailty conditions among older adults, facilitating early intervention. Goonawardene et al. [31] successfully combined sensor technology with machine learning for unobtrusive frailty detection. However, it is important to note that these advancements were made possible through studies largely reliant on data collected in a controlled clinical environment, underscoring the need for further exploration using real-world data. In a recent significant contribution, Kumar et al. [23] employed a tri-axial accelerometer sensor fixed in a t-shirt at the sternum to create digital biomarkers for frailty assessment. They focused on gait performance parameters, achieving an accuracy, sensitivity, specificity, and a *Receiver Operating Characteristic* (ROC) *Area Under the Curve* (AUC) of 0.78, 0.77, 0.80, and 0.84, respectively. This study underscores the potential of wearable technology and machine learning in providing an effective frailty assessment tool in real-world settings, a goal our current research also aims to accomplish.

In a previous study [32], we demonstrated that data collected with a small device worn at the wrist during a brief (60-m) supervised walk, processed with Machine Learning and Continuous Wavelet Transform, yielded satisfactory discrimination between robust and non-robust, where the latter category includes both frail and prefrail subjects. In the

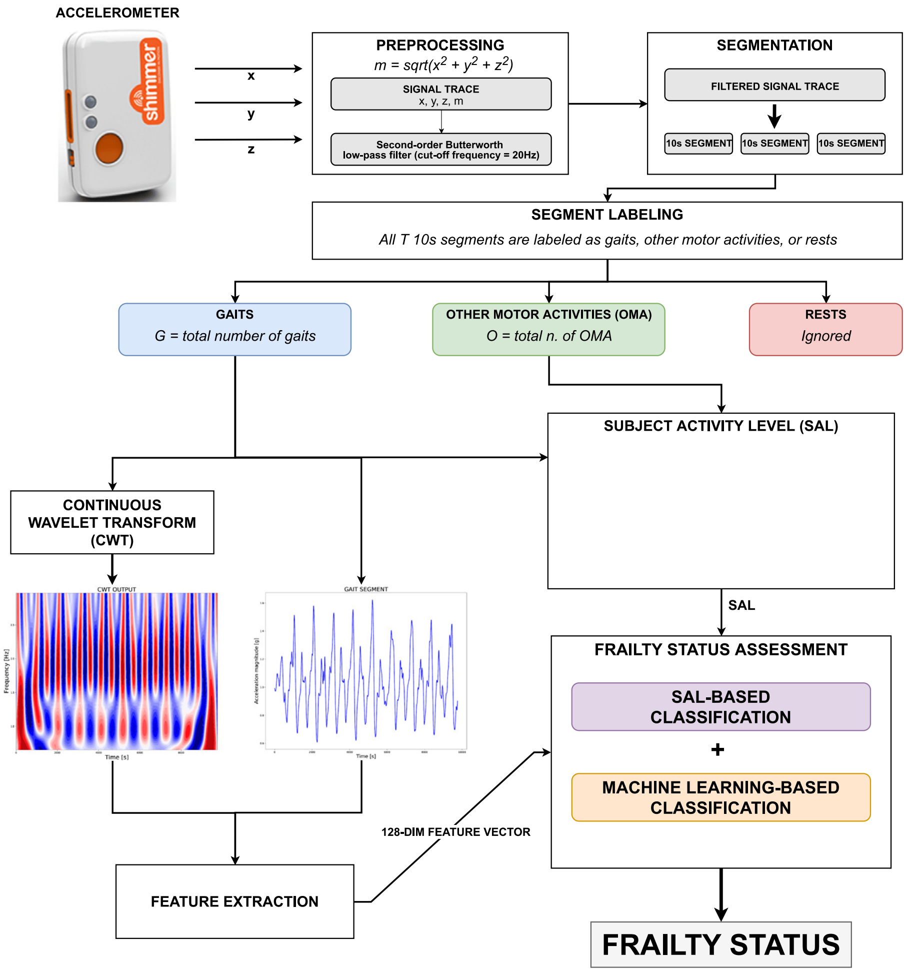


Fig. 1. Flowchart of the proposed method.

current study, we have extended data collection to a 24-h period, during which participants wore the wrist sensor in their homes, thus allowing for an ecologic assessment of frailty in a real-world setting. Leveraging on our initial experience, we applied an algorithm to identify gait instances and considered them as relevant to classify frailty status. However, we developed a new, comprehensive algorithm that considered the entire spectrum of motor activities, not limited to gait.

The manuscript is organized as follows: Section 2 presents the method we suggested for automatic frailty assessment based on wearable accelerometers. The experimental setup, design, and validation are discussed in Section 3. Finally, the results achieved and conclusions are presented and discussed in Sections 4 and 5.

2. The proposed method

Fig. 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-s 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). Additionally, we utilize a feature extraction phase that focuses on a specific subset of gait segments. This subset is determined by their average absolute acceleration variation (AAV) values, prioritizing those with the highest AAV value to ensure that the most informative segments are included in our machine learning model. Finally, the extracted features are fed into 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 non-robust (NR).

2.1. 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 highlighted by the authors in [33], 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.

2.2. Segmentation

To ensure uniformity of the input data across all subjects, we divide the signal trace into segments of fixed duration. 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.

We performed an analysis where the segment duration was varied, and the score of our system was computed for each window length. We found that a *duration* = 10 s yielded the highest classification performance compared to shorter or longer window durations. From now on, we will refer to the output of the segmentation phase with *10s segments*. In our study, each subject produced a total of 8640 segments per day, calculated by dividing the total number of seconds in 24 h (86400 s) by the length of each segment (10 s).

2.3. Segment labeling

Several researchers have stated the importance of activity levels in assessing different clinical conditions. We evaluate the subjects' activeness by exploring the nature of their 10 s segments. In particular, we apply a two-stage algorithm to analyze the content of each segment and label it as gait, other motor activity, or rest.

First, gait segments are automatically identified through the gait detection algorithm developed in [34] and adapted in [32]. 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. All those segments that remain unlabeled would be moved to the next stage. Next, 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. All the remaining segments are labeled as *rest segments* and discarded from the system. At this point, other motor activities and gait segments can be used to evaluate the subject's activity level.

2.4. 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 [35]. 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 h), 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 T be the total number of 10 s segments identified in a subject, and g, o, and r the number of subject's gaits, other motor activities, and rests, respectively. We will have the following:

$$G = \frac{g}{T}$$

$$O = \frac{o}{T}$$

$$R = \frac{r}{T}$$

where G, O, and R are the subject's gait level, other motor activities level, and rest level, respectively. It is worth noting that $G + O + R = 1$. The SAL is calculated as:

$$SAL = (1 - \alpha)O + \alpha G, \tag{1}$$

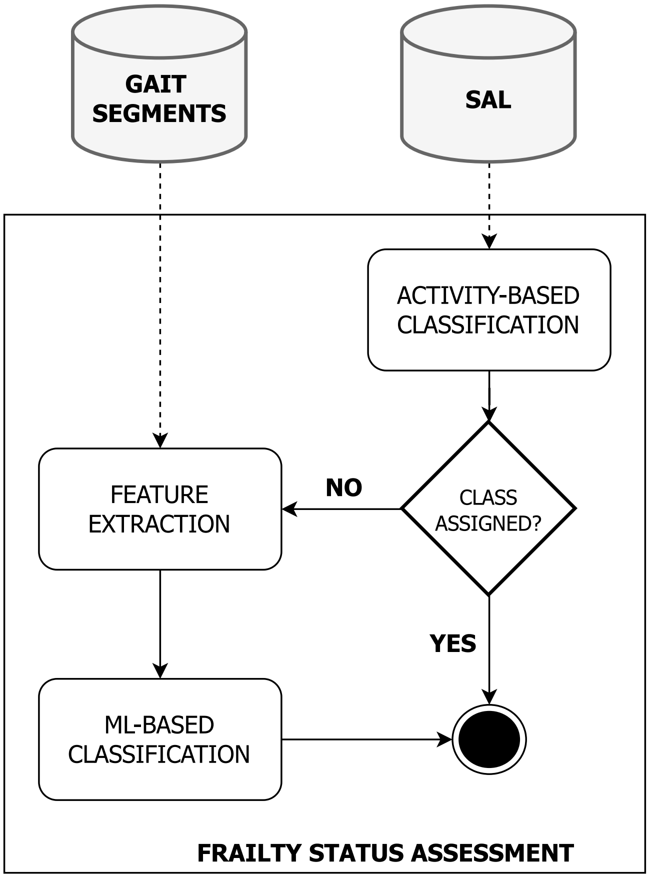


Fig. 2. Frailty status assessment.

where α is a parameter to weight G and O. In particular, we hypothesized that the importance of G and O 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 3.

2.5. 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 Fig. 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 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. The thresholds were established based on the percentile values of the activity levels distribution within the training set.

$$\begin{cases} SAL \leq AL_{NR} \rightarrow classifiedasNR \\ SAL \geq AL_R \rightarrow classifiedasR \\ otherwise \rightarrow ML - based\ classification \end{cases}$$

The ML-based classification module is a slightly modified version of what we already proposed in [32]. The walks performed by a subject during the day may exhibit diverse patterns, for instance due to variations in speed. This could potentially introduce noise into the training phase of the ML model. To mitigate this problem, we employ an initial filtering step where only the gaits with the highest energy levels are selected. Specifically, gait segments are sorted according to their average absolute acceleration variation (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.

Finally, a Machine Learning model is used to assess the frailty status. Let us define as *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 1.

Algorithm 1: ML-based classification algorithm

Result: Subject's frailty status (R or NR)
Let $G = \{g_1, \dots, g_M\}$ be the set of subject's gait instances;
foreach $g_i \in G$ **do**
  g_i is classified as R or NR;
end
Let $T = \{t_1, \dots, t_M\}$, 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, $k_{NR} = 0.6$;
threshold $\leftarrow \lceil M \cdot k_{NR} \rceil$;
if $T_{NR} \geq$ *threshold* **then**
  The subject is classified as NR;
else
  The subject is classified as R;
end

In the next Section, we provide more details on the estimation of the SAL thresholds, on the feature extraction procedure, as well as on the k_{NR} parameter mentioned above.

3. 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.

3.1. 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) [36], 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.

3.2. 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) [37];
3. low physical activity, defined thanks to the Physical Activity Questionnaire for the Elderly (PASE) [38];
4. slowness, defined by the speed measured over a distance of 4.5 m and normalized for height and gender;
5. weakness, meaning reduced handgrip 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.

Our 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 h, during which participants led their usual lives. Shimmer3 is a wearable device embedding a tri-axial accelerometer (STMicro LSM303DLHC) [39], which was used to collect acceleration samples at 102.4 Hz. At the end of the monitoring period, 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* [40].

3.3. Subject activity level

As mentioned in Section 2.4, we introduced a measure that quantifies the level of activity of the subject during the day: the SAL value. In addition, as showed in Eq. (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 4.

3.4. 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). These features are calculated on acceleration components and Wavelet coefficients. Furthermore, two quantities generally used in gait analysis and fall detection studies are computed: the *cadence*, defined 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 [34,41].

The importance of walk-related wavelet features for frailty status assessment has already been shown in [32]. 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 [42]. More specifically, CWT allows the analysis of local variations of power by decomposing a time series into different frequency components. A vector of statistics is computed from the output of the CWT and concatenated with the time-domain feature vector.

3.5. Evaluation 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 test set, while the data of other subjects were used as the training set. This methodology provides a robust measure of statistical significance by considering the variability in performance across different subsets of the data, ensuring reliability and real-world applicability. Besides, SAL values of the training set were used, at each iteration, to compute the AL_{NR} and AL_R thresholds introduced in Section 2.5. 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 4.1.

Five different Machine Learning models have been evaluated in the ML-based classification module: Random Forest, Gaussian Naive Bayes, Logistic Regression, Multilayer Perceptron, and Support Vector Machine. All the classifiers have been implemented by means of the Python module Scikit-learn [43]. 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. During our experiments, various k values were examined, and $k = 35$ was empirically determined to deliver the highest performance in terms of classification accuracy, sensitivity, specificity, and AUC, therefore ensuring optimal generalization and efficacy in our frailty detection model. The ML model was then trained using only the selected features of the training set and tested on the left-out subject.

In the context of tuning our ML models, it is pertinent to understand that this process involved adjusting the hyper-parameters of each model. For the Random Forest classifier, we focused on the number of trees in the forest, whereas for the Logistic Regression model, the tuning centered around the type of regularization (L1, L2, or ElasticNet) and the corresponding C parameter, controlling the regularization strength. Our strategy involved implementing the GridSearchCV method of the Python module Scikit-learn, which allowed us to explore the hyper-parameter space exhaustively. During each iteration of the Leave One Subject Out Cross Validation, this tuning was performed strictly using the training set data. This approach was intended to enhance the robustness and impartiality of our evaluations.

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 the 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.

4. Results and discussion

The results obtained from the experiments described in Section 3 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.

4.1. Subject activity level

As already described in Section 3.3, 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.

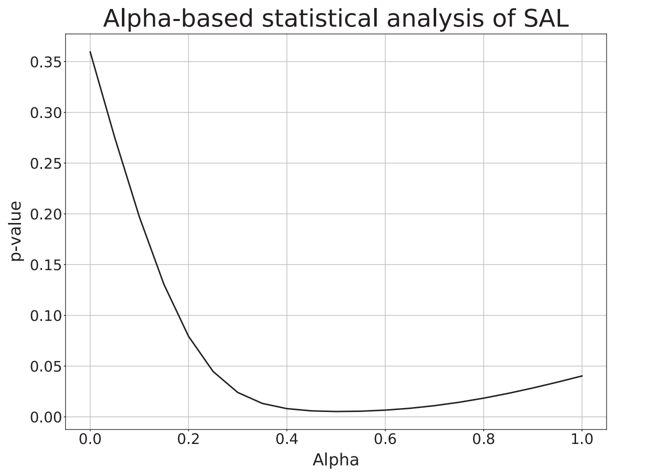
Fig. 3 summarizes the results of steps 1 and 2 of our α-based experiments, respectively. On the x-axis of the graph depicted in Fig. 3(a), we find the alpha parameter of the SAL, while on the y-axis, we have the p-value calculated by the statistical t-test R SAL vs. NR SAL. It is important to remember how a lower p-value corresponds to a higher statistical significance of the SAL in distinguishing between R and NR subjects. In Fig. 3(b), 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 – $SAL = (1 - \alpha)O + \alpha G$ – 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. It is important to consider 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 subjects, which in the case of our study, are captured by an almost equally balanced SAL.

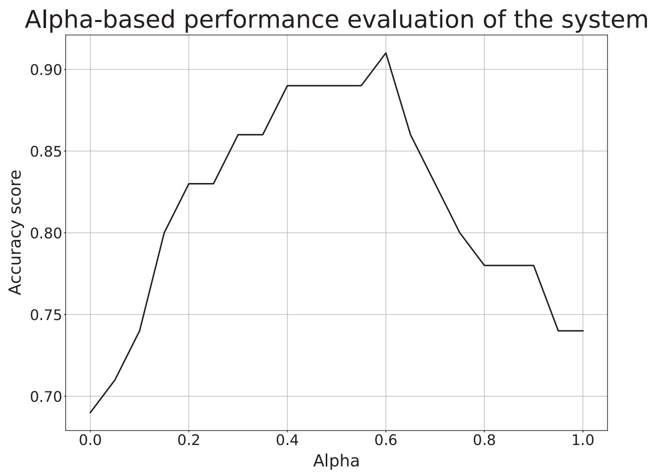
In light of what we obtained from our investigation, we assigned $\alpha = 0.6$, and the SAL formula became as follows:

$$SAL = 0.4O + 0.6G,$$

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. Fig. 4 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, we ensured the validity of the evaluation procedure.



(a)



(b)

Fig. 3. Results obtained in the α-based statistical analysis of SAL and in the α-based performance evaluation of the system.

Table 1
Average results of the frailty status assessment algorithm for each one of the tested models.

Classifier	Acc.	Sens.	Spec.	AUC
Gaussian Naive Bayes	0.91	0.94	0.88	0.91
Random Forest	0.77	0.68	0.88	0.78
Multilayer Perceptron	0.77	0.95	0.56	0.75
Logistic Regression	0.74	0.74	0.75	0.74
Support Vector Machine	0.69	0.68	0.69	0.69

4.2. 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 process described in Section 2. The frailty status assessment algorithm was adapted to the study context by tuning the parameters and thresholds defined in Section 3. Finally, we tested the entire system through a LOSO CV procedure. Table 1 summarizes the results obtained for all five classifiers tested.

The Gaussian Naive Bayes model performed substantially better than the other models, classifying subjects with an accuracy score of 91% (32 correctly classified participants out of 35) and an AUC of 0.91. In addition, the algorithm

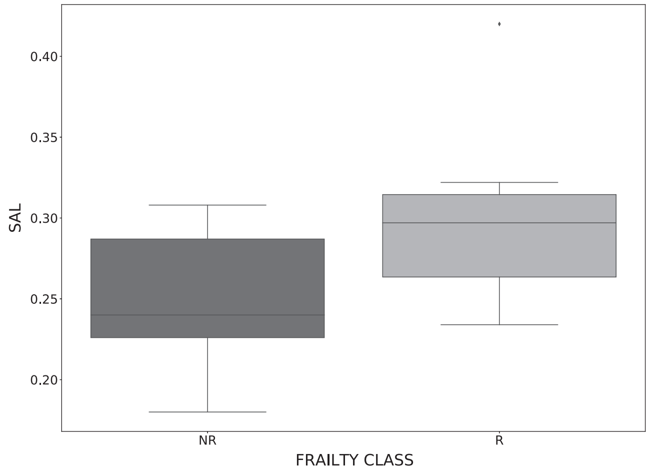


Fig. 4. Box plot representation of the SAL computed from the R and NR subjects.

recognized 18 out of 19 NR subjects (94% sensitivity score) and 14 out of 16 R subjects (88% specificity score). Random Forest and Multilayer Perceptron achieved the second-best accuracy of 77%. Nevertheless, Random Forest reached a low sensitivity score, a crucial indicator in healthcare-related applications (low sensitivity leads to false negatives, i.e., the system misses frail patients). On the other hand, Multilayer Perceptron achieved unbalanced scores. Logistic Regression and Support Vector Machine 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 [32]. 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 (specifically Gaussian Naive Bayes), may enable a practical approach to continuous frailty status monitoring. Second, this study 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 the Introduction, 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.

5. Conclusions

Wearable technology has gained a critical role in medical research, as it enables the study of health-related digital biomarkers in unsupervised, non-clinical environments. In this study, 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 five Machine Learning classifiers: Random Forest, Gaussian Naive Bayes, Logistic Regression, Multilayer Perceptron, and Support Vector Machine. 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.

The limited number of subjects involved in the experiments represents a limitation of this study. 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. 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.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Domenico Minici reports financial support was provided by Government of Italy Ministry of Education University and Research.

Data availability

The data that has been used is confidential.

Acknowledgments

This research was supported by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence) and by the PRA 2018_81 project "Wearable sensor systems: personalized analysis and data security in healthcare" funded by the University of Pisa.

References

[1] N.C. Hantke, C. Gould, Examining older adult cognitive status in the time of COVID-19, J. Am. Geriatr. Soc. (2020).
[2] O.S. Albahri, A. Zaidan, B. Zaidan, M. Hashim, A.S. Albahri, M. Alsalem, 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, J. Med. Syst. 42 (9) (2018) 1–47.
[3] T.H. Sanders, A. Devergnas, T. Wichmann, M.A. Clements, Remote smartphone monitoring for management of Parkinson's Disease, in: Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments, 2013, pp. 1–5.
[4] R. Gilpin, The Challenge of Global Capitalism, Princeton University Press, 2018.
[5] Mental health of older adults, 2017, https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults. (Accessed 13 July 2022).
[6] E. Dent, F.C. Martin, H. Bergman, J. Woo, R. Romero-Ortuno, J.D. Walston, Management of frailty: opportunities, challenges, and future directions, Lancet 394 (10206) (2019) 1376–1386.
[7] R.M. Collard, H. Boter, R.A. Schoevers, R.C. Oude Voshaar, Prevalence of frailty in community-dwelling older persons: a systematic review, J. Am. Geriatr. Soc. 60 (8) (2012) 1487–1492.
[8] R.J. Gobbens, K.G. Luijkx, M.T. Wijnen-Sponselee, J.M. Schols, Toward a conceptual definition of frail community dwelling older people, Nurs. Outlook 58 (2) (2010) 76–86.
[9] Q.-L. Xue, The frailty syndrome: definition and natural history, Clin. Geriatr. Med. 27 (1) (2011) 1–15.
[10] L.P. Fried, C.M. Tangen, J. Walston, A.B. Newman, C. Hirsch, J. Gottdiener, T. Seeman, R. Tracy, W.J. Kop, G. Burke, et al., Frailty in older adults: evidence for a phenotype, J. Gerontol. (A Biol. Sci. Med. Sci.) 56 (3) (2001) M146–M157.
[11] A. Mañas, B. del Pozo-Cruz, F.J. García-García, A. Guadalupe-Grau, I. Ara, Role of objectively measured sedentary behaviour in physical performance, frailty and mortality among older adults: A short systematic review, Eur. J. Sport Sci. 17 (7) (2017) 940–953.
[12] E. García-Esquinas, E. Andrade, D. Martínez-Gómez, F.F. Caballero, E. López-García, F. Rodríguez-Artalejo, Television viewing time as a risk factor for frailty and functional limitations in older adults: results from 2 European prospective cohorts, Int. J. Behav. Nutr. Phys. Activity 14 (1) (2017) 1–18.
[13] L. Dipietro, W.W. Campbell, D.M. Buchner, K.I. Erickson, K.E. Powell, B. Bloodgood, T. Hughes, K.R. Day, K.L. Piercy, A. Vaux-Bjerke, et al., Physical activity, injurious falls, and physical function in aging: an umbrella review, Med. Sci. Sports Exerc. 51 (6) (2019) 1303.
[14] D.S. Kehler, J.L. Hay, A.N. Stammers, N.C. Hamm, D.E. Kimber, A.S. Schultz, A. Szwajcer, R.C. Arora, N. Tangri, T.A. Duhamel, A systematic review of the association between sedentary behaviors with frailty, Exp. Geront. 114 (2018) 1–12.
[15] D.S. Kehler, O. Theou, The impact of physical activity and sedentary behaviors on frailty levels, Mech. Ageing Dev. 180 (2019) 29–41.
[16] C. Ziller, T. Braun, C. Thiel, Frailty phenotype prevalence in community-dwelling older adults according to physical activity assessment method, Clin. Interv. Aging 15 (2020) 343.
[17] A. de Carvalho Bastone, E. Ferriolli, C.P. Teixeira, J.M.D. Dias, R.C. Dias, Aerobic fitness and habitual physical activity in frail and nonfrail community-dwelling elderly, J. Phys. Activity Health 12 (9) (2015) 1304–1311.
[18] J. Blodgett, O. Theou, S. Kirkland, P. Andreou, K. Rockwood, The association between sedentary behaviour, moderate–vigorous physical activity and frailty in NHANES cohorts, Maturitas 80 (2) (2015) 187–191.
[19] L. Dasenbrock, A. Heinks, M. Schwenk, J. Bauer, Technology-based measurements for screening, monitoring and preventing frailty, Z. Gerontol. Geriatr. 49 (7) (2016) 581–595.
[20] L.H. Kikkert, M.H. De Groot, J.P. van Campen, J.H. Beijnen, T. Hortobágyi, N. Vuillerme, C.C. Lamoth, Gait dynamics to optimize fall risk assessment in geriatric patients admitted to an outpatient diagnostic clinic, PLoS One 12 (6) (2017) e0178615.
[21] S. Chen, T. Chen, H. Kishimoto, H. Yatsugi, S. Kumagai, Associations of objectively measured patterns of sedentary behavior and physical activity with frailty status screened by the frail scale in Japanese community-dwelling older adults, J. Sports Sci. Med. 19 (1) (2020) 166.
[22] M. Ritt, S. Schülein, H. Lubrich, L. Bollheimer, C. Sieber, K.-G. Gassmann, 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, J. Nutr. Health Aging 21 (3) (2017) 346–353.
[23] D. Pradeep Kumar, N. Toosizadeh, J. Mohler, H. Ehsani, C. Mannier, K. Laksari, Sensor-based characterization of daily walking: a new paradigm in pre-frailty/frailty assessment, BMC Geriatr. 20 (2020) 1–11.

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

[25] M. Huisingh-Scheetz, K. Wroblewski, M. Kocherginsky, E. Huang, W. Dale, L. Waite, L.P. Schumm, The relationship between physical activity and frailty among US older adults based on hourly accelerometry data, J. Gerontol. (A Biol. Sci. Med. Sci.) 73 (5) (2018) 622–629.

[26] N. Toosizadeh, B. Joseph, M.R. Heusser, T.O. Jokar, J. Mohler, H.A. Phelan, B. Najafi, Assessing upper-extremity motion: an innovative, objective method to identify frailty in older bed-bound trauma patients, J. Am. Coll. Surg. 223 (2) (2016) 240–248.

[27] A. Mulasso, P.R. Brustio, A. Rainoldi, G. Zia, L. Feletti, A. N'dja, S. Del Signore, E. Poggiogalle, F. Luisi, L.M. Donini, A comparison between an ICT tool and a traditional physical measure for frailty evaluation in older adults, BMC Geriatr. 19 (1) (2019) 1–7.

[28] H. Lee, B. Joseph, A. Enriquez, B. Najafi, Toward using a smartwatch to monitor frailty in a hospital setting: using a single wrist-wearable sensor to assess frailty in bedbound inpatients, Gerontology 64 (4) (2018) 389–400.

[29] D. Ravi, C. Wong, B. Lo, G.-Z. Yang, A deep learning approach to on-node sensor data analytics for mobile or wearable devices, IEEE J. Biomed. Health Inf. 21 (1) (2016) 56–64.

[30] A. Tarekegn, F. Ricceri, G. Costa, E. Ferracin, M. Giacobini, et al., Predictive modeling for frailty conditions in elderly people: machine learning approaches, JMIR Med. Inf. 8 (6) (2020) e16678.

[31] N. Goonawardene, H.-P. Tan, L.B. Tan, Unobtrusive detection of frailty in older adults, in: Human Aspects of IT for the Aged Population. Applications in Health, Assistance, and Entertainment: 4th International Conference, ITAP 2018, Held As Part of HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part II 4, Springer, 2018, pp. 290–302.

[32] 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 J. Biomed. Health Inf. 26 (3) (2021) 1013–1022.

[33] E.K. Antonsson, R.W. Mann, The frequency content of gait, J. Biomech. 18 (1) (1985) 39–47.

[34] G. Cola, A. Vecchio, M. Avvenuti, Improving the performance of fall detection systems through walk recognition, J. Ambient Intell. Humaniz. Comput. (ISSN: 1868-5137) 5 (6) (2014) 843–855, http://dx.doi.org/10.1007/s12652-014-0235-x.

[35] A.P. Tolley, K.A. Ramsey, A.G. Rojer, E.M. Reijnierse, A.B. Maier, Objectively measured physical activity is associated with frailty in community-dwelling older adults: A systematic review, J. Clin. Epidemiol. 137 (2021) 218–230.

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

[37] J.G. Orme, J. Reis, E.J. Herz, Factorial and discriminant validity of the center for epidemiological studies depression (CES-D) scale, J. Clin. Psychol. 42 (1) (1986) 28–33.

[38] A.J. Schuit, E.G. Schouten, K.R. Westerterp, W.H. Saris, Validity of the Physical Activity Scale for the Elderly (PASE): according to energy expenditure assessed by the doubly labeled water method, J. Clin. Epidemiol. 50 (5) (1997) 541–546.

[39] Shimmer, http://www.shimmersensing.com.

[40] B. Granger, F. Pérez, Jupyter: Thinking and storytelling with code and data, 2021, Authorea Preprints.

[41] G. Cola, M. Avvenuti, F. Musso, A. Vecchio, Personalized gait detection using a wrist-worn accelerometer, in: 2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks, BSN, 2017, pp. 173–177, http://dx.doi.org/10.1109/BSN.2017.7936035.

[42] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. Inform. Theory 36 (5) (1990) 961–1005.

[43] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.