

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Dataset of acceleration signals recorded while performing activities of daily living



Pau Climent-Pérez^{a,*}, Ángela M. Muñoz-Antón^a, Angelica Poli^b, Susanna Spinsante^b, Francisco Florez-Revuelta^a

^a Department of Computing Technology, University of Alicante, Alicante 03690, Spain ^b Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, Ancona 60131, Italy

ARTICLE INFO

Article history: Received 21 December 2021 Revised 17 January 2022 Accepted 25 January 2022 Available online 3 February 2022

Keywords:

Accelerometry data Active and assisted living Human action recognition Wearable device Wrist-worn device

ABSTRACT

Several research studies have investigated the human activity recognition (HAR) domain to detect and recognise patterns of daily human activities. However, the accurate and automatic assessment of activities of daily living (ADLs) through machine learning algorithms is still a challenge, especially due to limited availability of realistic datasets to train and test such algorithms. The dataset contains data from 52 participants in total (26 women, and 26 men). The data for these participants was collected in two phases: 33 participants initially, and 19 further participants later on. Participants performed up to 5 repetitions of 24 different ADLs. Firstly, we provide an annotated description of the dataset collected by wearing a wrist-worn measurement device, Empatica E4. Secondly, we describe the methodology of the data collection and the real context in which participants performed the selected activities. Finally, we present some examples of recent and relevant target applications where our dataset can be used, namely lifelogging, behavioural analysis and measurement device evaluation. The authors consider the dissemination of this dataset can highly benefit the research community, and specially those involved in the recognition of ADLs, and/or in the removal of cues that reveal identity.

* Corresponding author.

E-mail addresses: pau.climent@ua.es, pcliment@dtic.ua.es (P. Climent-Pérez).

Social media: 🔰 @unQpassava (P. Climent-Pérez), 🔰 @AngelicaPoli4 (A. Poli), 🔰 @suspin76 (S. Spinsante), 🔰 @fflorezrevuelta (F. Florez-Revuelta)

https://doi.org/10.1016/j.dib.2022.107896

2352-3409/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Specifications Table

SubjectSignal processingSpecific subject areaAccelerometer data for Active and assisted living: technologies for extended autonomy of older people, recognition of activities of daily living (ADLs), human action recognition (HAR), de-identification of subject-dependent traits (gender, age, dominant hand).Type of dataTable, FigureHow the data were acquiredEach subject performed the activities wearing the Empatica E4 bracelet on their dominant hand. Each subject took 30-40 minutes to complete the activities including repetitions.Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLS), users.csv with participant subject IDs, ages, and provided gender, and a data.zip file containing all comma-separated values (.csv) files named as				
extended autonomy of older people, recognition of activities of daily living (ADLs), human action recognition (HAR), de-identification of subject-dependent traits (gender, age, dominant hand).Type of dataTable, FigureHow the data were acquiredEach subject performed the activities wearing the Empatica E4 bracelet on their dominant hand. Each subject took 30-40 minutes to complete the activities including repetitions.Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip	Subject	Signal processing		
living (ADLs), human action recognition (HAR), de-identification of subject-dependent traits (gender, age, dominant hand).Type of dataTable, FigureHow the data were acquiredEach subject performed the activities wearing the Empatica E4 bracelet on their dominant hand. Each subject took 30-40 minutes to complete the activities including repetitions.Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLS), users.csv with participant subject IDs, ages, and provided gender, and a data.zip	Specific subject area	8 8		
Type of datasubject-dependent traits (gender, age, dominant hand).Type of dataTable, FigureHow the data were acquiredEach subject performed the activities wearing the Empatica E4 bracelet on their dominant hand. Each subject took 30-40 minutes to complete the activities including repetitions.Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904 (v10, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLS), users.csv with participant subject IDs, ages, and provided gender, and a data.zip				
Type of dataTable, FigureHow the data were acquiredTable, FigureData formatEach subject performed the activities wearing the Empatica E4 bracelet on their dominant hand. Each subject took 30-40 minutes to complete the activities including repetitions.Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904 (v10, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip				
How the data were acquiredEach subject performed the activities wearing the Empatica E4 bracelet on their dominant hand. Each subject took 30-40 minutes to complete the activities including repetitions.Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v10, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip				
Data formaton their dominant hand. Each subject took 30-40 minutes to complete the activities including repetitions. Raw FilteredDescription of data collectionS2 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v10, 33 initial participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip	51			
Data formatthe activities including repetitions.Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/4750904 (v1.0, 33 initial participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip	How the data were acquired			
Data formatRaw FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip				
FilteredDescription of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		the activities including repetitions.		
Description of data collection52 participants (33+19) were recorded in total (26 men, and 26 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling.Data source location• Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: SpainData accessibilityRaw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/4750904 (v1.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and adta.zip	Data format	Raw		
 women). All participants were asked about their dominant hand, gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling. Data source location Institution: University of Alicante (offsite data sourcing) City/Town/Region: Alicante province (several locations) Country: Spain Data accessibility Raw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/4750955 (v2.0, all 52 participants) Instructions for accessing these data: Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip 		Filtered		
gender, and age. They then performed 24 different ADLs, up to 5 times each. A video (non-disclosed), was used for labelling. Data source location Institution: University of Alicante (offsite data sourcing) City/Town/Region: Alicante province (several locations) Country: Spain Data accessibility Raw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip	Description of data collection	52 participants (33+19) were recorded in total (26 men, and 26		
each. A video (non-disclosed), was used for labelling. Data source location Institution: University of Alicante (offsite data sourcing) City/Town/Region: Alicante province (several locations) Country: Spain Data accessibility Raw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/4750904 (v1.0, 33 initial participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		women). All participants were asked about their dominant hand,		
Data source location • Institution: University of Alicante (offsite data sourcing) • City/Town/Region: Alicante province (several locations) • Country: Spain Data accessibility Raw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with		gender, and age. They then performed 24 different ADLs, up to 5 times		
 City/Town/Region: Alicante province (several locations) Country: Spain Country: Spain Raw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip 		each. A video (non-disclosed), was used for labelling.		
 Country: Spain Data accessibility Raw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip 	Data source location	 Institution: University of Alicante (offsite data sourcing) 		
Data accessibility Raw data is provided on the Zenodo repository at: Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		 City/Town/Region: Alicante province (several locations) 		
Repository name: Zenodo (zenodo.org) Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v10, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		• Country: Spain		
Data identification number: 4750904, and 5785955 Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip	Data accessibility	Raw data is provided on the Zenodo repository at:		
Direct URL to data: https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		Repository name: Zenodo (zenodo.org)		
 https://zenodo.org/record/4750904 (v1.0, 33 initial participants) https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip 		Data identification number: 4750904, and 5785955		
 https://zenodo.org/record/5785955 (v2.0, all 52 participants) Instructions for accessing these data: Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip 				
Instructions for accessing these data: • Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		https://zenodo.org/record/4750904 (v1.0, 33 initial participants)		
• Access the v2.0 link above, there should be three files: ADLs.csv with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		https://zenodo.org/record/5785955 (v2.0, all 52 participants)		
with the names of all 24 activities (ADLs), users.csv with participant subject IDs, ages, and provided gender, and a data.zip		Instructions for accessing these data:		
participant subject IDs, ages, and provided gender, and a data.zip				
		with the names of all 24 activities (ADLs), users.csv with		
file containing all comma-separated values (.csv) files named as		participant subject IDs, ages, and provided gender, and a data.zip		
		file containing all comma-separated values $(.csv)$ files named as		
described in 'Data format' above.		described in 'Data format' above.		
Filtered data can be generated via a provided script [1].		Filtered data can be generated via a provided script [1].		
Related research article A. Poli, A. Muñoz-Antón, S. Spinsante, F. Florez-Revuelta, Balancing	Related research article	A. Poli, A. Muñoz-Antón, S. Spinsante, F. Florez-Revuelta, Balancing		
activity recognition and privacy preservation with multi-objective				
evolutionary algorithm, in: Proc. 7th EAI International Conference on		evolutionary algorithm, in: Proc. 7th EAI International Conference on		
Smart Objects and Technologies for Social Good, 2021, pp. 1–15		Smart Objects and Technologies for Social Good, 2021, pp. 1–15		
https://doi.org/10.1007/978-3-030-91421-9_1		https://doi.org/10.1007/978-3-030-91421-9_1		

Value of the Data

- This set of data is useful for the training and testing of novel human action recognition methods from accelerometer data. Furthermore, systems built upon action recognition, can be used for automated *lifelogging* (for self-observation, and reflection), and long-term behaviour analysis: e.g. decline in autonomy, or variations in performance, in an active and assisted living (AAL) context.
- Existing datasets (see Table 1 below) tend to have a low number of participants; these tend to be young (very often 20–30 y/o, rarely older than 50); imbalanced in gender (usually more men); are mostly recorded in lab conditions, rather than real-life scenarios using everyday use articles (combs, irons, crockery and cutlery, etc.); and classes (labels) consist of *motion primitives* (standing, sitting, walk, lie, bend), rather than complex, *domestic* activities of daily living (washing dishes, ironing, dusting, etc.).

- This data can benefit researchers in the area of ADL recognition, specially those methods where identity privacy is to be preserved (i.e. recognition as an *optimisation* problem: maximising recognition of activities, while minimising identification of individual traits). It is also beneficial to the society at large, due to the current trend of ageing societies in developed nations, with increasing pressure on social and care services for the older population.
- The data can be used *simply* for recognition of human activities from accelerometer data, or, as proposed by the authors, to also minimise leakage of identity. Furthermore, if interested in gender or regression/estimation of age, the data could be used to infer either or both of these traits; for instance, for re-identification of individuals based on their characteristic motion patterns.

1. Data Description

Existing similar datasets (accelerometer-based) are shown in Table 1. It can be observed that the proposed dataset has one of the highest number of participants (52), and the highest number of class labels (24 different ADLs). 'MPs' stands for 'motion primitives' *only*, i.e. simpler movements or poses (e.g. standing, walking, sitting), whereas 'ADLs' encompass MPs *and* more complex *household activities* (washing hands, ironing, etc.).

As stated above, in the proposed dataset, a total of 52 participants (26 men, and 26 women) were recorded carrying out 24 different activities. Fig. 1 shows the distribution of participants according to their age (binned in age groups: 20s to 80s), and gender (two labels: 'men' and 'women'). Furthermore, Table 2 shows the 24 activities that are part of the dataset, along with the labels that were used for them, as well as a short description about the definition of each activity (what each particular ADL entailed, according to the researchers).

Regarding the selection of activities (class labels) for the proposed dataset, the NTU RGB+D 120 dataset [19] was chosen as a basis, since it is one of the largest datasets used for activity recognition in the field of computer vision (CV). It has 120 labels, including, among others, activities of daily living (ADLs) of which 24 were selected for the creation of the proposed dataset. Apart from ADLs the criteria also included activities which had a significant motion of the hands. Furthermore, the activities were chosen to be of several aspects of life, since decreased autonomy, caused by mild to moderate cognitive impairment of older people is usually assessed by

Dataset	Year	Participants	Activities
WISDM [2]	2010	36	6 MPs
WISDM 2.0/ActiTracker [3]	2012	59	6 MPs
UCI HAR [4]	2012	30	6 MPs
Casale et al. [5]	2012	10-20	7 MPs
ADL [6]	2013	16	14 ADLs
Barshan et al. [7]	2014	8	19 MPs (sport)
Mobifall [8]	2014	24	9 MPs + 4 falls
SAR [9]	2014	10	7 MPs
mHealth [10]	2015	10	12 MPs (sport)
Stisen et al. [11]	2015	9	6 MPs
JSI+FoS [12]	2016	15	10 MPs
ADLs dataset [13]	2017	-	14 ADLs
ASTRI [14]	2019	11	5 MPs
Intelligent Fall [15]	2019	6/11	16 ADLs + 5 falls
IM-WSHA [16]	2020	10	11 ADLs
Fioretti et al. [17]	2021	36	6 ADLs
Proposed: PAAL ADL v1.0 [18]	2021	33	24 ADLs
Proposed: PAAL ADL v2.0	2021	52	24 ADLs

 Table 1

 'Accelerometer data'-based datasets in the literature.

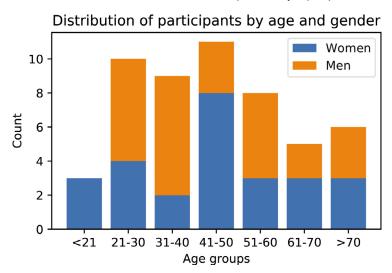


Fig. 1. Histogram showing the distribution of participants among different age groups and genders.

Table 2

List of activities included in the dataset. There are 24 different ADLs, and each participant provides up to 5 repetitions each. The activities can be divided into 6 broad categories: eating, and drinking (1–4); hygiene/grooming (5, 6); dressing and undressing (7–12); miscellaneous and communication (13–18); basic health indicators (19–21); and house cleaning (22–24).

Index	Activity (label)	Description
1	drink_water	Drink (once) from a glass, cup, or bottle.
2	eat_meal	Perform the gesture of eating, using a fork, a spoon, or the hands.
3	open_a_bottle	Open a bottle (uncap it) once.
4	open_a_box	Open a food container (e.g. Tupperware), once.
5	brush_teeth	Brush teeth for approximately 20 seconds.
6	brush_hair	Brush hair during 10 seconds (using a comb, or the hands).
7	take_off_a_jacket	Take off a jacket by undoing the buttons or zip (if zipped or buttoned).
8	put_on_a_jacket	Put on a jacket and optionally do the buttons or zip.
9	put_on_a_shoe	Put on a shoe, doing the laces, zip, etc. (if <i>available</i>)
10	take_off_a_shoe	Take off a shoe, by optionally undoing the laces/zip.
11	put_on_glasses	Put on (sun)glasses once.
12	take_off_glasses	Take off (sun)glasses once.
13	sit_down	Sit down on an (arm)chair/sofa/high stool, once.
14	stand_up	Stand up once.
15	writing	Write (by hand) for 15 to 20 seconds.
16	phone_call	Pick up the (mobile) phone once (bring to ear).
17	type_on_a_keyboard	Type on a computer/laptop keyboard for 15-20 seconds.
18	salute (wave hand)	Wave the hand once.
19	sneeze_cough	Sneeze or cough once.
20	blow_nose	Blow one's nose once.
21	washing_hands	Wash hands: apply soap, rub hands together, and rinse.
22	dusting	Dust a surface with a rag/cloth for some time (15-20 s).
23	ironing	Iron (a garment) for 15-20 s.
24	washing_dishes	Scrub/scour a plate, cup/glass, or fork/knife/spoon; and rinse.

human experts rating the performance of activities within different aspects of their daily routines (hygiene, home and hobbies, shopping, etc.) [20].

Because it was important to capture the most natural performance of activities from participants, these were recorded performing all activities in a single, or several, sessions in their



Fig. 2. ADLs performed in real-life conditions: a) ironing clothes, b) washing dishes.

workplace or their homes, using their everyday objects to perform the activities (e.g. iron and board, cutlery and crockery, kitchen sink, etc), as shown in Fig. 2 a) and b).

In each session, the bracelet was initialised, and the user started performing the activities with no particular order. Using the video (non-disclosed, for privacy), the researchers then created ground truth files, that were used to split the acceleration file from the Empatica E4 (ACC.csv), into several smaller .csv files that follow a naming convention: <activity_name>_<subject_id>_<repetition>.csv, for example, for the first repetition of activity phone call by individual number 34, the file would be named phone_call_34_0.csv. Given all possible combinations, of participant, repetition, and activity, it results in $52 \times 5 \times 24 = 6,240$ accelerometer data files in .csv format. However, due to the range of possible repetitions per participant, the total number of files provided is 6,072. Fig. 3 presents a colour-coded matrix showing the number of repetitions (colour) for each participant (ID in the x-axis) and activity (label in the y-axis). As can be observed, other than a few exceptions, most participants were recorded between 3 and 5 times per activity. This can be seen more clearly in Fig. 4, where the mean and standard deviation of the number of repetitions per activity (across all participants) are shown; and also Fig. 5 where the mean and standard deviation of repetitions per participant (across all activities) is plotted. This latter figure shows, however, that participant 17, for instance, has the lowest mean of repetitions, which has some implications for the usage of this dataset, specially if using leave-one-actor-out (LOAO) cross-validation: if used for validation or testing (when data is divided into folds) results for this particular participant might be unreliable (either too good, or too bad). This need to be taken into account in any possible division (splitting) of data.

To end the analysis of the accelerometer data files provided, Fig. 6 shows the mean and standard deviation for activity duration, according to the different types (activity labels). As shown in Table 2, in some activities, the participants were asked to perform a task during a certain amount of time, whereas other activities had more diverse lengths due to differences in partici-

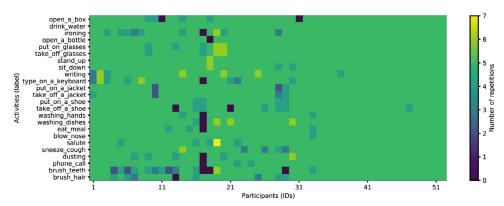


Fig. 3. A matrix plot showing the number of **repetitions per activity** and participant (IDs). As explained, participants provided up to 5 repetitions of each of the 24 activities considered.

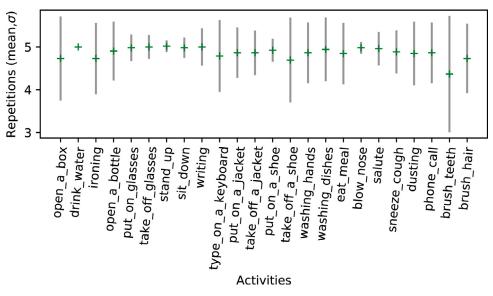


Fig. 4. Mean and standard deviation plot for activity repetitions across all participants in the dataset. It can be observed that most activities are between 3 and 5 repetitions on average.

pant performance. Duration of an activity in seconds is derived from the file length (number of samples), and divided by the sampling frequency (32 Hz for the device's accelerometer sensor).

Apart from the 6,072 accelerometer data files in .csv format, two further files are included in the dataset: one with all activities, with their indices (ADLs.csv), as shown in Table 2 (excluding descriptions, that is); and another one (users.csv) with participant (*user*) IDs, gender, and age (used as ground truth information for age regression and gender classification).

An example of the time variation of the acceleration signals along the three directions (x, y, and z) collected during the *washing dishes* activity is shown in Fig. 7.

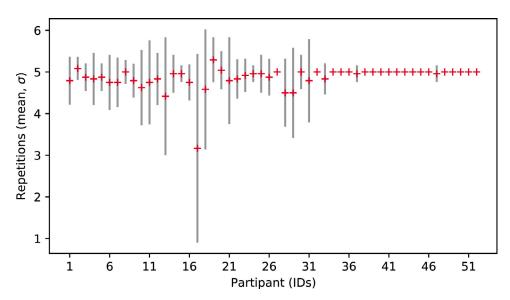


Fig. 5. Mean and standard deviation plot for repetitions per participant across all activities in the dataset. Please observe, except participant 17, all other participants provided sufficient repetitions.

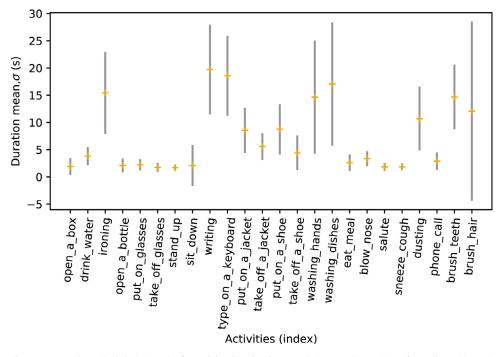


Fig. 6. Mean and standard deviation plot for activity duration in seconds (across all repetitions from all participants). As expected, some activities take longer to perform than others, but this should be taken into account for the design of a classification model.

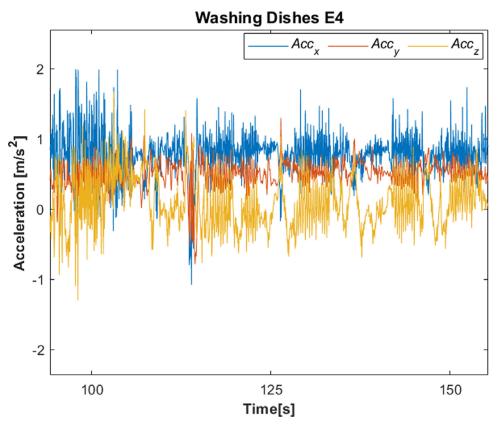


Fig. 7. Example of acceleration signals along x, y, and z directions collected by wearing E4 during the washing dishes activity.

2. Experimental Design, Materials and Methods

2.1. Measurement device: The Empatica E4

The Empatica E4 is a wrist-worn top-quality sensor device considered as a Class IIa Medical Device according to CE Crt. No. 1876/MDD (93/42/EEC Directive). Empatica E4 device measures the acceleration data, as well as other physiological parameters, namely the Blood Volume Pulse (BVP), from which the Heart Rate Variability (HRV) and the Inter Beat Interval (IBI) are derived as well, skin temperature (SKT) and also changes in certain electrical properties of the skin such as the Electrodermal Activity (EDA). For the creation of our dataset, among the several measurements recorded by the Empatica E4, only the 3-axis acceleration signal was considered, since it provides information better suited for activity recognition [6]. In particular, Empatica E4 is equipped with an accelerometer sensor (sampling frequency: 32 Hz), that measures the continuous gravitational force (i.e., g) exerted along the three spatial directions (i.e. x, y and z axis). By default, the range of scale is set to $\pm 2g$, but $\pm 4g$ or $\pm 8g$ can be set by requesting a custom firmware. A summary of the technical specifications of the accelerometer sensor is detailed in Table 3.

For the collection of the dataset, the accelerometer sensor is configured to measure acceleration in the range $[-2g, 2g] m/s^2$. Therefore, according to the measurement range, for analytic purposes, the conversion factor between raw acceleration samples and true values is equal to

-	
Specification	Value
Sampling Frequency (f_s) Resolution Range Time needed for automatic calibration	32 Hz 0.015g ±2g 15 s

Technical specifications of the accelerometer sensor (Empatica E4).

g/64 (where $g = 9.81 m/s^2$), that is, a sample value of 64 corresponds to 1g. Regarding the sensor's calibration, E4 calibrates automatically during the initial 15 s of each session. Finally, the device offers two operating modes: a *live* mode, in which data is streamed via Bluetooth to a mobile phone for visualisation, with the in-app option to also store the data as it is received; or, alternatively, a *recording* mode, in which the data is stored directly on the device's internal memory. Upon connection via USB to a host computer, the data capture sessions can then be synchronised (copied) from the device.

2.2. Collection procedure

Table 3

With the good results observed in a previous work [17], and expanding on the initial data collection of 33 participants (v1.0), during the months of October and November 2021, an additional 19 participants were recorded using an Empatica E4 bracelet (using the *recording*, offline mode), following the same procedure in both cases: participants were asked about their dominant hand, their age and gender, and provided with a participant ID number. This information was stored in a file as described. Next, they were told to perform up to 5 repetitions of a set of 24 activities (either in one or several sessions). Most recordings took place in the homes or workplaces of individuals, thus collecting the activities as normally performed in the subject's environment. No restrictions or instructions were given, except the use of the dominant hand to carry out the activities, as well as to remind them that each repetition should be performed *independently* from each other (e.g. if washing hands, assume hands are dry at the start of the repetition, even if they are still wet from previous repetition). Additionally, the activities were performed with the help of daily life objects, which are usually not provided when recording in lab conditions.

Individuals were recorded performing these activities while wearing the bracelet on their reported dominant hand: the additional video footage (non-disclosed) was used to assist the researchers in the task of ground truth labelling. Synchronisation of video and bracelet sensor information was performed using either the 'tagging' mechanism of the Empatica E4 (pressing of the bracelet's button makes an LED illuminate for 1 s and saves a *timestamp* to a tags.csv file on the device); and/or the bracelet's LED status change, i.e. registering the video frame in which the LED goes solid red during start-up, indicating the start of the sensor capture (first *timestamp* registered on the device).

Using a video media player software providing millisecond accuracy of frames¹, the LED status was tracked, and annotated, either from the status change at initialisation (LED change from blue to red), or from the moment of 'tapping' (LED change from off to red). This timestamp annotation, in the video, was then used to synchronise both data streams (video and accelerometer data file from the bracelet). This is possible because at initialisation, the Empatica E4 bracelet saves the timestamp when the device starts recording as the first line of the internally stored accelerometer data file (ACC.csv). Furthermore, if synchronised via the 'tapping' of the device's button, the internal tags.csv file will store the timestamp of each time the button was 'tapped'. From that point onwards, using the video player screen, all activity start and end times-

¹ For instance, for participants 34–52, the software package 'Avidemux' was used for this purpose.

tamps were annotated in a text file. These text files were then processed by a Python script (not included) to split each session's ACC.csv file into several, smaller, labelled .csv files (each of the 6,072 data files provided).

Ethics Statements

As per European Regulation 2016/679, i.e. the General Data Protection Regulation (GDPR), a written informed consent was obtained from all the participants prior to starting the data collection, in order to obtain the permission for processing personal data. All participants were provided with information about the study and the type of data collected prior to any data capture. They were given the opportunity to continue or withdraw from the study at any point without further questioning. This process was carried out following the ethics protocols established by the authors' institutions. Furthermore, the data is anonymised, and identities of the participants are not revealed, nor can be obtained from the published data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT Author Statement

Pau Climent-Pérez: Data curation, Investigation, Formal analysis, Methodology, Software, Writing – original draft; **Ángela M. Muñoz-Antón:** Data curation, Investigation, Formal analysis, Methodology, Software, Writing – original draft; **Angelica Poli:** Data curation, Investigation, Formal analysis, Methodology, Software, Writing – original draft; **Susanna Spinsante:** Conceptualization, Supervision, Methodology, Software, Resources, Funding acquisition, Writing – review & editing; **Francisco Florez-Revuelta:** Conceptualization, Supervision, Methodology, Software, Resources, Funding acquisition, Writing – review & editing.

Acknowledgments

The support of the More Years Better Lives JPI, the Italian Ministero dell'Istruzione, Università e Ricerca (CUP: I36G17000380001) and the Spanish Agencia Estatal de Investigación (grant no: PCIN-2017-114), for this research activity carried out within the project PAAL – Privacy-Aware and Acceptable Lifelogging services for older and frail people, (JPI MYBL award number: PAAL JTC2017) is gratefully acknowledged.

The authors would also like to acknowledge all anonymous participants who took part in this study for their valuable time and contribution to the collection of this dataset.

References

- P. Climent-Pérez, Ángela María Muñoz-Antón, A. Poli, S. Spinsante, F. Florez-Revuelta, HAR filtered features generator code, 2022, doi:10.5281/zenodo.5849344.
- [2] J.R. Kwapisz, G.M. Weiss, S.A. Moore, Activity recognition using cell phone accelerometers, ACM SigKDD Explorat. Newsletter 12 (2) (2011) 74–82.
- [3] G.M. Weiss, J. Lockhart, The impact of personalization on smartphone-based activity recognition, in: Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence, 2012, pp. 98–104.
- [4] J.-L. Reyes-Ortiz, D. Anguita, A. Ghio, X. Parra, Human activity recognition using smartphones data set, UCI Machine Learning Repository; University of California, Irvine, School of Information and Computer Sciences: Irvine, CA, USA (2012).
- [5] P. Casale, O. Pujol, P. Radeva, Personalization and user verification in wearable systems using biometric walking patterns, Pers. Ubiquit. Comput. 16 (2012) 1–18.

- [6] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, R. Zaccaria, Analysis of human behavior recognition algorithms based on acceleration data, in: 2013 IEEE International Conference on Robotics and Automation, IEEE, 2013, pp. 1602–1607.
- [7] B. Barshan, M.C. Yüksek, Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units, Comput. J. 57 (2014) 1649–1667.
- [8] G. Vavoulas, M. Pediaditis, C. Chatzaki, E.G. Spanakis, M. Tsiknakis, The mobifall dataset: Fall detection and classification with a smartphone, Int. J. Monitor. Surveillance Technol. Res. (IJMSTR) 2 (1) (2014) 44–56.
- [9] M. Shoaib, S. Bosch, O.D. Incel, H. Scholten, P.J. Havinga, Fusion of smartphone motion sensors for physical activity recognition, Sensors 14 (6) (2014) 10146–10176.
- [10] O. Banos, C. Villalonga, R. García, A. Saez, M. Damas, J. Holgado-Terriza, S. Lee, H. Pomares, I. Rojas, Design, implementation and validation of a novel open framework for agile development of mobile health applications, BioMed. Eng. OnLine 14 (2015) S6.
- [11] A. Stisen, H. Blunck, S. Bhattacharya, T.S. Prentow, M.B. Kjærgaard, A. Dey, T. Sonne, M.M. Jensen, Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition, in: SenSys '15: 13th ACM Conference on Embedded Networked Sensor Systems, Association for Computing Machinery, 2015, pp. 127–140.
- [12] M. Gjoreski, H. Gjoreski, M. Luštrek, M. Gams, How accurately can your wrist device recognize daily activities and detect falls? Sensors 16 (6) (2016), doi:10.3390/s16060800.
- [13] W. Gomaa, R. Elbasiony, S. Ashry, Adl classification based on autocorrelation function of inertial signals, in: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 2017, pp. 833–837, doi:10.1109/ ICMLA.2017.00-53.
- [14] J. Lu, K.-y. Tong, Robust single accelerometer-based activity recognition using modified recurrence plot, IEEE Sensor. J. 19 (15) (2019) 6317–6324.
- [15] L. Chen, R. Li, H. Zhang, L. Tian, N. Chen, Intelligent fall detection method based on accelerometer data from a wrist-worn smart watch, Measurement 140 (2019) 215–226, doi:10.1016/j.measurement.2019.03.079.
- [16] A. Jalal, M.A.K. Quaid, S.B.u.d. Tahir, K. Kim, A study of accelerometer and gyroscope measurements in physical life-log activities detection systems, Sensors 20 (22) (2020), doi:10.3390/s20226670.
- [17] S. Fioretti, M. Olivastrelli, A. Poli, S. Spinsante, A. Strazza, Adls detection with a wrist-worn accelerometer in uncontrolled conditions, in: P. Perego, N. TaheriNejad, M. Caon (Eds.), Wearables in Healthcare, Springer International Publishing, Cham, 2021, pp. 197–208.
- [18] A. Poli, A. Muñoz-Anton, S. Spinsante, F. Florez-Revuelta, Balancing activity recognition and privacy preservation with multi-objective evolutionary algorithm, in: in Proc. 7th EAI International Conference on Smart Objects and Technologies for Social Good, 2021, pp. 1–15.
- [19] J. Liu, A. Shahroudy, M. Perez, G. Wang, L-Y. Duan, A.C. Kot, Ntu rgb+d 120: a large-scale benchmark for 3d human activity understanding, IEEE Trans. Pattern Anal. Mach. Intell. 42 (10) (2020) 2684–2701, doi:10.1109/TPAMI.2019. 2916873.
- [20] J.C. Morris, The Clinical Dementia Rating (CDR), Neurology 43 (11) (1993) 2412–2412–a, doi:10.1212/WNL43.11. 2412-a.