

Wearable Sensors and MET Features: The Importance of Preprocessing and Understanding the Data

Maria Bermudez-Edo^{1,2} , Alberto Durán-López^{1,2} , Daniel Bolaños-Martínez^{1,2*} , Jose Luis Garrido^{1,2} , Maria Jose Rodriguez-Fortiz^{1,2} 

¹ Department of Software Engineering, Computer Science School, University of Granada, Granada (Spain)

² Research Centre for Information and Communication Technologies (CITIC-UGR), University of Granada, Granada (Spain)

* Corresponding author: danibolanos@ugr.es

Received 26 May 2025 | Accepted 16 May 2026 | Published 28 August 2026



ABSTRACT

Discerning user activities with wearable devices is important to monitor people's behavior and even for developing personalized strategies to prevent some diseases, which has significant implications for healthcare systems. The metabolic equivalent of the task (MET), a measure of the energy cost of physical activities, is a good indicator to discern different activities. However, there is no established formula for calculating the MET. Some works have used several features to approximate the MET, such as the filtered acceleration (ACCfil) or the percentage of the heart rate reserve (%HRR). However, none of them has studied the importance of each of these features in the accuracy of activity classification. This study investigates the impact of different preprocessing techniques on activity recognition accuracy, such as filtering the acceleration to obtain the feature ACCfil. To that end, we performed several experiments with different features and machine learning algorithms to detect the activities. Our results indicate that incorporating features such as ACCfil, %HRR, and Ratio of Unfiltered to Filtered acceleration (RUF) improves activity recognition accuracy in some cases up to 166.67%. These features are particularly useful for challenging activities such as stair climbing or household activities. These findings, with their direct implications for developing more accurate activity recognition models with wearable devices, underscore the importance of preprocessing techniques and feature integration.

KEYWORDS

Feature Selection, Human Activity Recognition, Metabolic Equivalent of the Task, Preprocessing, Sensors.

DOI: 10.9781/ijimai.2026.2257
eLocator: e2257

I. INTRODUCTION

ASSESSING physical activity and quantifying the frequency and intensity of different activities are important across various fields, for instance, in developing strategies aimed at preventing various chronic diseases [1]–[3]. In this regard, the metabolic equivalent of the task (MET) offers a reliable approach to classify physical effort by expressing activity intensity as a multiple of the resting metabolic rate (RMR) [4], [5]. This measurement can be facilitated by utilizing accelerometers and PhotoPlethysmoGraphy (PPG) sensors integrated into wearable devices [6]–[9]. However, it is important to note that sensors do not directly measure the MET; instead, MET is an approximate value that attempts to correlate the activity with its Energy Expenditure (EE). While some wearables provide MET data, they generally do not disclose the specific formula used for the calculation, and the accuracy of the MET is dubious [10], [11].

Certain studies have attempted to estimate the MET by establishing correlations between accelerometer data and known MET values for

specific activities (e.g., a MET of 3 for walking, as established in the MET reference tables [12]). Some of these studies emphasize the significance of isolating the two components of acceleration: the gravitational acceleration in lower frequencies and the body movement acceleration in higher frequencies [13]. Key factors used in MET measurement include: filtered acceleration obtained by removing gravitational effects; the Ratio of Unfiltered to Filtered acceleration (RUF); the Sum Vector Magnitude (SVMg) of the filtered acceleration ($SVM_{g_{fil}}$); and the percentage Heart Rate Reserve (%HRR) [14]. Although MET is directly correlated to human activity recognition (HAR), the use of MET or MET-related characteristics, such as $SVM_{g_{fil}}$ or RUF, are not standard features used in HAR studies [13] due to the difficulty of determining the correct MET automatically. Furthermore, the impact of using these features compared to traditional approaches in HAR still needs to be explored.

Hence, we propose a comparative analysis of activity detection models developed solely based on acceleration data, which is the most common way to apply machine learning (ML) to HAR [13], against

models incorporating one or more of the features mentioned above, correlated with the MET. To calculate these features is necessary a preprocessing step, which involves filtering the acceleration signals and computing the %HRR.

Our study specifically evaluates the effectiveness of integrating features such as $SVMg_{fil}$, %HRR, and RUF in enhancing the accuracy of activity detection, particularly for challenging activities such as household chores and stair climbing/descending. These activities are known to present difficulties in accurate differentiation when considering only acceleration data [15].

Hence, our contribution can be summarized as:

- Highlighting the importance of preprocessing the sensor signals, especially filtering the gravitational and body acceleration components in HAR studies.
- Comparison of using different MET related features, namely $SVMg_{fil}$, %HRR, and RUF in HAR studies vs not using them.
- Our results concluded that using MET-related features considerably increase the accuracy of HAR studies.

The content of the paper is as follows. Section II discusses the present research works carried out in the domain of HAR and MET. Section III discusses the proposed methodology, experimental setup and analysis. The results are given in Section IV followed by discussion in Section V and conclusion in Section VI.

II. BACKGROUND AND RELATED WORKS

A. Background: Measuring MET With Wearables

EE could be a good discriminator in activities, as different activities need different energy. However, the gold standards to measure EE are the Doubly Labeled Water (DLW) and Indirect Calorimetry (IC) techniques [16]. DLW includes administering doubly labeled water and monitoring saliva, urine, or blood; hence, an expert must collect and analyze the samples. IC measures pulmonary gas exchanges and needs a specific device to measure oxygen consumption and carbon dioxide production, which is uncommon in popular wearables. To solve this problem of devices and human resources availability, the use of MET as an approximation of the EE has emerged. One MET represents the oxygen consumption at rest [16], approximately $3.5 : ml, O_2 / kg / min, (1.2, kcal / min$ for a 70kg person) [12]. Actual energy cost differs for each person depending on body mass, adiposity, age, gender, and environmental conditions [16]. Several activities have been assessed regarding their relative oxygen uptake or intensity as multiples of 1 MET, creating tables with activities and their METs [12]. However, these are fixed tables that cannot be automatically measured in real daily life activities. Wearable devices frequently serve as convenient and relatively cost-effective alternatives for estimating EE, primarily through collecting Heart Rate (HR) and accelerometer data. However, no perfect algorithm exists to estimate EE [17]. Some wearables measure the EE, but while the accuracy in HR and other indicators is high, the accuracy in EE is low [10], [11], [18], [19]. Moreover, the accuracy is even lower in some groups, such as those with darker skin tones or males with greater body mass index [19]. Some studies use rule-based algorithms to estimate the MET based on features such as $SVMg_{fil}$, %HRR, or RUF [14].

B. Related Works: Using MET in HAR

A relevant example of the calculation of the MET with wearables is the study carried out by [20], in which they developed a rule-based decision tree algorithm that can be used in wearable health monitoring devices to estimate MET based on physical activity intensity data. The proposed algorithm uses filtered synthetic acceleration (ACCfil) and RUF ratio. Similarly, Nakanishi et al. [14] enhance the rule-based

calculation of MET with the %HRR. Studies have recently used these calculations to estimate the EE [21], [22]. These studies calculate the MET in two steps. First, they compute the features, ACCfil, RUF, etc., and then apply the rule to estimate the MET. They calculate these features for each activity using the whole activity to estimate MET without windowing. Then, the activities were distinguished using a rule-based algorithm between low, medium, and high MET activities. To estimate the MET as EE and create the rule-based formula, the most used algorithms are regression algorithms [14], [20], [22].

HAR studies, not using rule-based calculation of MET, have used several ML algorithms; some of the most commonly used include Logistic Regression (LR), KNN (K-Nearest Neighbors), families of the Decision Tree (DT), and Gaussian Naïve Bayes (NB) [13], [23]. For example, [24], [25] uses Random Forest (RF), KNN, and Support Vector Machines (SVN, linear, and RBF). [26] uses J48, RF, NB, SVM, Neural Networks, and KNN. They also compared the Empatica wristband with an Arduino-based accelerometer, showing that Empatica was superior in performance and accuracy.

However, only a few HAR studies have used the MET or related features. For example, Barut et al. [27] used one ML model to distinguish activities with accelerometer, gyroscope, and magnetometer data. Then, they improved the model with a calculated MET based on accelerometer counts. The climbing and descending stairs activities were the most difficult to distinguish. They performed no feature extraction and just fed the raw data into an LSTM model. Similarly, [28] used the SVMg of the filtered accelerometer, among other features, to create a model for HAR. They positioned the accelerometer in the lower leg and back. They used random forests to classify between walking, sitting, lying, and running, which are easier to distinguish than climbing/descending stairs or household chores. [29] filtered the accelerometer signal and used the SVMg of the filtered acceleration, among other features, to detect the activities of nose/head scratching, smoking, drinking, eating, and using the phone. However, these studies used only some MET-related features directly in the HAR task, not all of them. Even some works tried to filter the accelerometer and showed that the filtered signal did not improve the accuracy but degraded it [24]. They used two low-pass and two high-pass filters at two different frequencies directly over the three dimensions of the accelerometer (x, y, z). They did not use the $SVMg_{fil}$ or the RUF.

Recent trends focus on Deep Learning (DL) architectures, generative models for data synthesis [30], and Large Language Models (LLMs) to handle data complexity [31]–[34]. While effective, these models often operate as "black boxes" and require substantial resources. Other emerging lines of research focus on specific constraints, such as feature optimization strategies [35], LoRaWAN connectivity for long-range networks [36], multi-environment data fusion [37], hybrid vision-based systems for fall detection [38], or energy-aware computing [39]. Despite this breadth, most approaches, even recent ones leveraging diffusion models or feature subsets [30], [35], rely solely on accelerometer data, leaving a specific gap in leveraging physiological domain knowledge directly within the classification pipeline. Table I describes an extended comparative analysis of related works in HAR, highlighting the research gap in MET-based feature engineering addressed by this study. Our study distinguishes itself by proposing a preprocessing approach that explicitly separates gravitational and body acceleration to derive MET-based features ($SVMg_{fil}$, RUF, %HRR). This approach prioritizes physiological interpretability and accuracy in complex activities, contrasting with the predominantly data-driven or connectivity-focused approaches found in the current state of the art.

TABLE I. COMPARATIVE ANALYSIS OF RELATED WORKS IN HAR, HIGHLIGHTING THE RESEARCH GAP IN MET-BASED FEATURE ENGINEERING ADDRESSED BY THIS STUDY

Ref. - Authors (Year)	Main Focus / Methodology	Identified Limitation / Gap	Our Contribution
[14] - Nakanishi et al. (2018)	Rule-based MET estimation enhanced with %HRR features	Classifies activities into broad intensity categories (Low/ Mid/High) using rules, lacking specific activity identification.	Physiological Feature Engineering: We propose a specific preprocessing pipeline that separates gravitational from body acceleration to derive MET-related features (SVMgfil, RUF, %HRR).
[20] - Ohkawara et al. (2011)	Rule-based decision tree for MET	Relies on a fixed rule-based decision tree structure, which	
[21] - Koizumi et al. (2023)	Estimation of EE	Focuses on regression algorithms for EE,	
[22] - Chen et al. (2023)	Sedentary behavior analysis and	Analyzes general sedentary behavior versus active states;	Advancement over MET Works: Unlike [27] which relied on raw data, or [28] which required intrusive sensor placement, we achieve high accuracy using a single wrist device. Furthermore, we refute claims that filtering degrades accuracy [24] by proving the efficacy of .
[24] - Erdaş et al. (2016)	Feature integration for accelerometer-based HAR	Concluded filtering degraded accuracy (applied to axes); lacked analysis of Vector Magnitude.	
[27] - Barut et al. (2020)	LSTM model + calculated MET counts	Fed raw data to LSTM; struggled to distinguish climbing/descending stairs effectively.	
[28] - Stewart et al. (2018)	Random Forest with SVMg of filtered acceleration	Required intrusive sensors (leg/back); focused primarily on simple activities (walking, lying).	Improved Discrimination: While existing works focus on generic DL models [32]-[34] or network constraints [36], we demonstrate that incorporating interpretable physiological metrics drastically improves accuracy for complex activities of daily living (ADLs) (e.g., household chores, stairs) without relying on black-box models.
[29] - Lu et al. (2020)	Single sensor HAR for specific hand habits	Limited scope (smoking, eating); did not integrate full physiological set (RUF, %HRR).	
[30] - Garcia-Moreno et al. (2025)	Time-series data generation with diffusion models for triaxial data	Relies solely on accelerometers; lacks feature selection (consequently missing MET-related features).	
[31] - Ferrara (2024)	LLMs & Generative AI for behavioral modeling	High computational cost; focus on high-level language/behavior, less on sensor signal physics.	
[32] - Gupta et al. (2022)	Narrative review of AI frameworks in HAR	Broad overview of general AI methods; lacks specific focus on domain-knowledge MET features.	
[33] - Kumar et al. (2024)	DL review & methodologies	Emphasis on complex DL architectures (black-box); interpretability of physical features is limited.	
[34] - Zhang et al. (2022)	DL advances with wearable sensors	Highlights DL trends but overlooks the value of signal filtering for metabolic estimation.	
[35] - Liandana et al. (2024)	Single triaxial accelerometer HAR using feature subset combination	Relies solely on kinematic features; lacks physiological data (HR) and MET estimation for intensity discrimination.	
[36] - Obiri & Van Laerhoven (2024)	LoRaWAN-integrated sensor networks survey	Focuses on IoT connectivity/transmission layers, not on signal preprocessing for activity intensity.	
[37] - Li et al. (2023)	Multi-environment sensor data fusion	Relies on external environmental sensors; lacks physiological data (HR) and does not employ MET-based features.	
[38] - Liu et al. (2025)	Hybrid Fall Detection (YOLO + ST-GCN)	Relies on visual data (cameras), implying privacy issues and fixed setups (non-wearable).	
[39] - Contoli et al. (2024)	Energy-aware HAR for battery efficiency	Prioritizes device battery optimization; overlooks the use of MET or metabolic features for activity discrimination.	

TABLE II: COMPARISON OF EXISTING PUBLIC DATASETS VERSUS THE DATASET COLLECTED IN THIS STUDY. NOTE: ACC = ACCELEROMETER, GYRO = GYROSCOPE, MAG = MAGNETOMETER, TEMP = TEMPERATURE, EDA = ELECTRODERMAL ACTIVITY.

Dataset	Device Model (Position)	Sensors	Key Activities	Gap for this Study
WISDM [40]	LG G Watch (Wrist) Nexus 5/5X (Pocket)	Acc, Gyro	Walking, Jogging, Stairs, Sitting, Standing	Lacks physiological data (HR) necessary for %HRR; limited to basic locomotive activities (no chores).
UCI HAR [41]	Samsung Galaxy S II (Waist)	Acc, Gyro	Walking, Upstairs, Downstairs, Sitting, Standing, Lying	Waist placement is not comparable to wrist-wearables; lacks HR sensor; no household activities.
UniMiB SHAR [42]	Galaxy Nexus I9250 (Pocket)	Acc	9 ADLs (Walk, Run, Jump, Stairs), 8 types of Falls	Pocket placement; relies solely on Accelerometer (no Gyro or HR); primarily focused on fall detection.
MobiAct [43]	Samsung Galaxy S3 (Pocket)	Acc, Gyro, Orientation	Walk, Jog, Jump, Stairs, Sit, Car step-in/out, Falls	Pocket placement; lacks physiological data (HR) required for metabolic analysis.
KU-HAR [44]	Various Smartphones (Waist)	Acc, Gyro	Stand, Sit, Talk, Lay, Pick, Jump, Push-up, Sit-up, Walk, Run, Stairs	Waist placement; lacks HR data; includes fitness activities (push-ups) but lacks specific household chores.
WEEE [45]	Empatica E4, Apple Watch, Fitbit (Wrist), Earbuds	Acc, Gyro, PPG, EDA, Temp	Resting (Sit/Stand), Cycling, Running	Contains HR but lacks activity diversity (only cardio/rest); focuses on EE regression rather than activity classification.
CAPTURE-24 [46]	Axivity AX3 (Wrist)	Acc	In-the-wild (Sleep, Indoor, Chores, Work, Walk, Sports)	Lacks HR sensor (AX3 is Acc only), preventing %HRR calculation; uncontrolled protocol makes MET rule validation difficult.
Our Dataset	Empatica E4 (Wrist)	Acc, PPG (HR)	Stairs (Up/Down) + Household (Dishes, Sweeping, Drawing) + Lying/Sitting/Walking	Specifically designed to combine wrist kinematics with HR to validate MET features (, RUF, %HRR) on hard-to-discern activities.

III. METHODS

A. Experiment Description: Participants and Tasks

There are several public datasets with accelerometer data and labeled activities. Popular examples include WISDM [40] and UCI HAR [41]. However, UCI HAR relies on waist-mounted sensors, while UniMiB SHAR [42] and MobiAct [43] collect data from smartphones in pockets. These placements differ significantly from wrist-worn wearables. Furthermore, datasets like KU-HAR [44] lack the physiological data

(PPG/Heart Rate) necessary for metabolic features like %HRR, and household chores. While the WEEE dataset [45] includes multimodal sensors (Acc + HR), it focuses on basic locomotive and cardio activities (cycling, running), neglecting complex household chores. Conversely, large-scale datasets like CAPTURE-24 [46] offer wrist-worn data using the Axivity AX3 (accelerometer-only), preventing %HRR calculation, and their "in-the-wild" protocol lacks the controlled segmentation required to validate specific MET-based rules. Table II describes the comparison between existing public datasets and the dataset collected

in this study. As detailed in the Table II, a distinct gap remains for a dataset that combines wrist-worn kinematics with physiological data to discern high-intensity activities like stair climbing from complex, energetically variable household chores. Hence, we create our own dataset and experiments that we discuss in this section.

We have performed an experiment, in which seven volunteers participated, within an age bracket of 16-53, performing activities of daily living, mainly household activities, resting activities, walking and climbing, and descending stairs. In particular, they performed ten activities, starting from the quiet ones and finished with the most active, in this order: lying in bed, sitting in a chair, drawing, walking at a regular pace, brisk walking, climbing stairs, descending stairs, sweeping the floor, washing dishes, and strolling. All participants wore a wearable, the Empatica E4 wristband, on the non-dominant wrist. All activities were longer than 3 minutes, and the first activity (lying on the bed) lasted for at least 7 minutes to capture the heart rate at rest, which is necessary to calculate the %HRR.

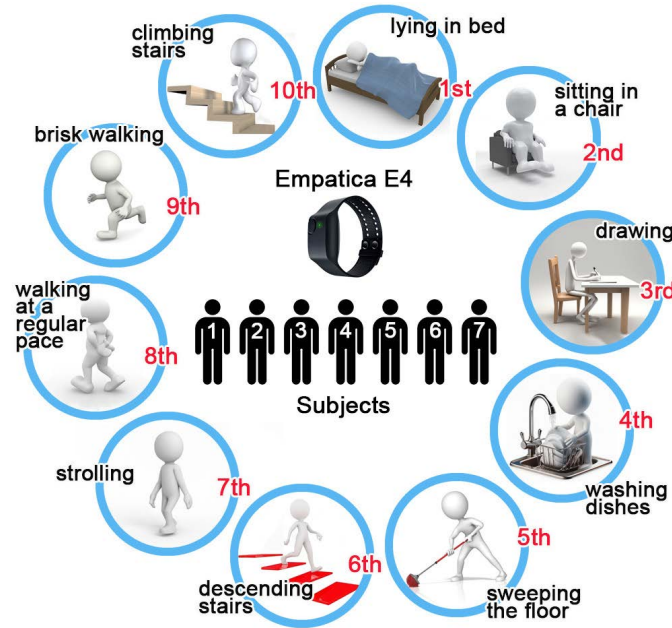


Fig. 1. Activities performed by volunteers during the experiment.

B. Materials

The Empatica E4 features four physical sensors: a PPG, an electrodermal activity sensor, a peripheral thermometer, and a triaxial accelerometer. The Empatica calculates values from the PPG, including the blood volume pulse, which represents the highpass filtered PPG, as well as the HR and interbeat intervals measured in milliseconds. Our study will focus on the accelerometer signal, a standard component of most wristbands and smartwatches widely used for activity recognition. Additionally, we will incorporate HR into our analysis, as it is an important factor in determining the MET. For the calculations, we used a laptop MacBookPro, with a chip M1 and 16GB. We create the pipeline with Python and use the libraries Scikit-learn with numpy and pandas.

C. Machine Learning Pipeline

We follow the pipeline illustrated in Fig. 2. In the first step of data acquisition, as the sampling rate of Empatica is 32Hz, it can only capture frequencies lower than 16Hz and thus produces one sample every 1/32 s, approximately 31.25 ms. Hence, we do not need to apply any lowpass filter to the accelerometer signal to reduce the noise of the high frequencies. If we apply any lowpass filter, we could miss

important information about body movements, which are typically contained within frequency components below 20 Hz (for example, in gait, 99% of the energy is contained below 15 Hz) [47]. Hence, we take the three components of the acceleration (x, y, z) and the HR directly from the Empatica raw data.

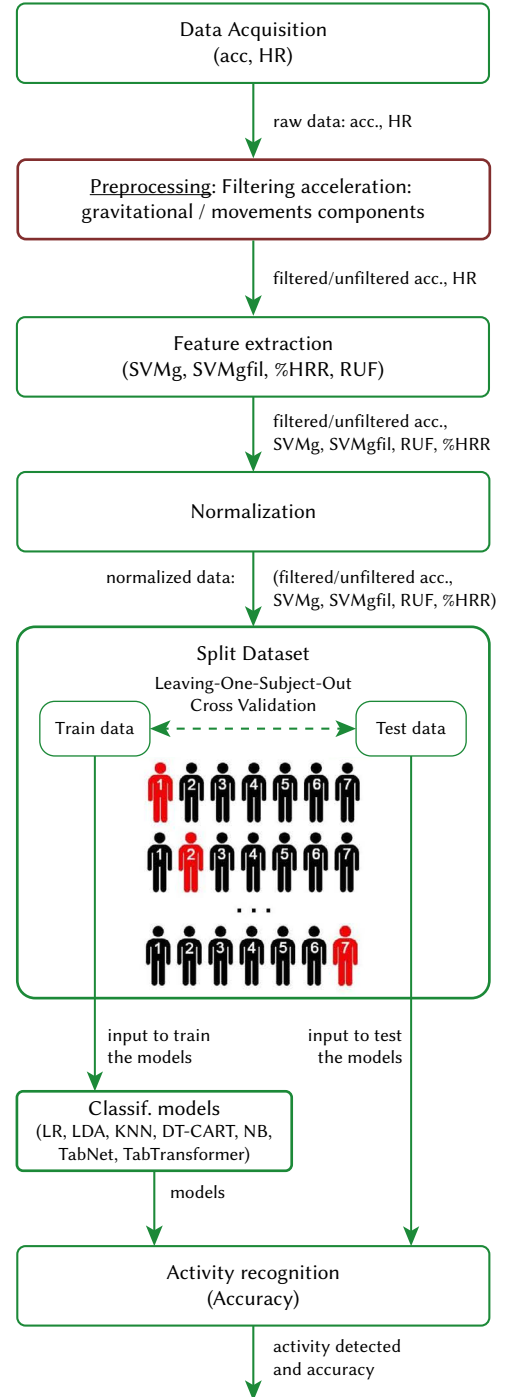


Fig. 2. Pipeline.

Then, in the second step of the pipeline, the preprocessing, we performed the proposed preprocessing of the data. We separate the body acceleration from the gravitational acceleration by applying a highpass filter, particularly a 4th-order butterfly highpass filter, with a cut-off frequency of 3 Hz, as in previous works [29]. The high frequencies of this separation are the filtered acceleration in each of the components of the acceleration, x_{fil} , y_{fil} , z_{fil} . The filtered acceleration

represents the body's movement acceleration after removing the gravitational component. In contrast, the raw acceleration (x, y, z) includes both the gravitational and movement acceleration.

In the third step of the pipeline, we calculate all the features that the literature has used in approximating the MET by rules. Specifically, we have followed the approach of [14] that combines all the features used by others. In particular, we extracted the features: x, y, z, and the SVMg as in Equation (1), xfil, yfil, zfil, and then with these values we calculate the SVMg of filtered acceleration ($SVMg_{fil}$) as in Equation (2), and the ration between SVMg and $SVMg_{fil}$ (RUF) as in Equation (3). We also calculate the %HRR with the Karkoven formula as in Equation (4), where the heart rate during activity (HRact) represents the actual HR during activities, while the heart rate at rest (HRrest) is the mean value of heart rate at rest. We consider rest as lying down for at least 7 minutes. We calculate the maximum heart rate (HRmax) as in Equation (5).

$$SVMg = \sqrt{x^2 + y^2 + z^2} \quad (1)$$

$$SVMg_{fil} = \sqrt{x_{fil}^2 + y_{fil}^2 + z_{fil}^2} \quad (2)$$

$$RUF = \frac{SVMg}{SVMg_{fil}} \quad (3)$$

$$\%HRR = \left(\frac{HR - HR_{rest}}{HR_{max} - HR_{rest}} \right) \times 100\% \quad (4)$$

$$HR_{max} = 220 - Age \quad (5)$$

The SVM ($SVMg$ and $SVMg_{fil}$) reduces the complexity of the data by collapsing three-dimensional acceleration data into a single value. This makes it easier to analyze and interpret the data, particularly in applications where the overall magnitude of movement is more important than the direction of the movement, which is often indicative of specific activities (e.g., walking, running, sitting). This can lead to more accurate and robust activity classification models. SVM captures the overall activity intensity more effectively than analyzing individual axis values separately, reducing the individual noise. The RUF can provide insights into the intensity of body movements relative to the gravitational component. Higher ratios might indicate more dynamic activities (where the movement component of the acceleration is more prominent), while lower ratios might suggest more static postures (where the gravitational component of the acceleration is more prominent). HRR normalizes the heart rate data relative to each individual's resting heart rate. This normalization helps compare heart rate data across different individuals more effectively and detect high- or low-intensity activity for each individual.

We do not apply any windowing as we aim to understand the importance of the features used, not the windowing used. Furthermore, each activity could have its particular window size to capture the activity efficiently [29], which is out of the scope of our study.

The fourth step of our pipeline is important for ensuring the reliability of our ML models. We normalize the data, ensuring that all features have the same range of values. This normalization, performed using a standard scaler, provides a mean of zero and a standard deviation of 1, thereby moving all values between the range [-1, 1]. This step is essential to prevent any bias in our models.

In the fifth step, we use the leaving-one-subject-out cross-validation (LOSO-CV), as it is the most realistic, as the model is usually created over a training subject and then used in real-time with the new buyers of the device or the app. It is also the most restrictive, as none of the data and subjects' idiosyncrasy will be mixed in both training and test. Generally, with this cross-validation, the results yield worse but more realistic values.

In step six, for each validation, we train the model with different ML algorithms commonly used in HAR. In particular, we use LR, Linear Discrimination Analysis (LDA), K-NN, Decision tree Classification And Regression Tree (DT-CART), and NB, implemented via the library scikit Learn¹. We also explored two DL algorithms based on attention mechanisms and Transformer architectures, which are TabNet [48] and TabTransformer [49], implemented via the pytorch-tabnet² and tab-transformer-pytorch³ libraries respectively. For all of them, we used the default parameters of the libraries, and we did not perform any hyperparameter tuning to avoid complexity and to maintain control over the features used, thereby ensuring a clear assessment of the contribution of each feature ($SVMg_{fil}$, RUF and %HRR).

Once the model is created, in the seventh step of the pipeline, we validate it with the test data and calculate its accuracy. We use the accuracy metric because it is widely used in HAR.

Then, we performed an ablation test, in which we repeated steps five and six with different inputs. First, we used only the raw data acceleration features, and then, in each iteration, we used a different set of features incorporating $SVMg_{fil}$, RUF, and %HRR.

IV. RESULTS

We begin by analyzing the correlation matrix to understand the influence of each feature on activity detection. Fig. 3 shows the correlation between the different variables, and Table III value details. This figure corresponds to one of the subjects but is similar to the rest. In this subject, the correlation between $SVMg_{fil}$ and the activity is almost 75%, while the correlation between the SVMg is almost 20% in all subjects. We could also see that the heart rate and the HRR have a correlation above 77%. However, some wearables do not have an HR sensor, and others do not provide continuous heart rate. Furthermore, this correlation depends highly on the subject, as it is shown in Table III. We can appreciate that subject-specific correlations for SVMg range from 2.0% (Subject 5) to 49.1% (Subject 3; mean 18.3%), whereas those for $SVMg_{fil}$ span 48.7% (Subject 1) to 74.9% (Subject 7; mean 63.4%).

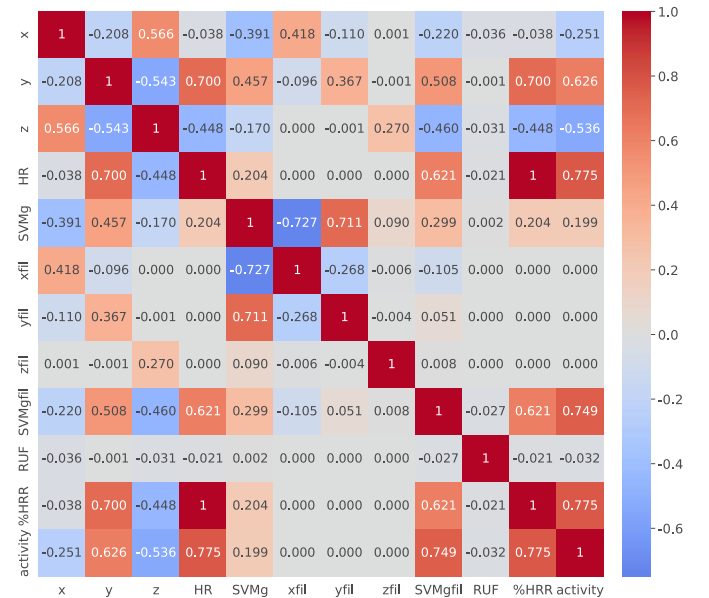


Fig. 3. Correlation Matrix of one of the subjects.

¹ <https://scikit-learn.org/>

² <https://pypi.org/project/pytorch-tabnet/>

³ <https://pypi.org/project/tab-transformer-pytorch/>

TABLE III. CORRELATION BETWEEN THE DIFFERENT VARIABLES AND THE ACTIVITIES FOR ALL OF THE SUBJECTS

	Subject 1	Subject 2	Subject 3	Activity Subject 4	Subject 5	Subject 6	Subject 7
x	-0.24989	-0.20073	0.04789	0.48842	0.66845	-0.32621	-0.25143
y	0.63821	0.64902	0.77660	0.69722	-0.46474	-0.84114	0.62631
z	-0.23703	-0.56945	-0.53217	-0.58636	-0.69077	-0.58687	-0.53592
HR	0.79810	0.67843	0.77200	0.91921	0.84761	0.76485	0.77491
SVMg	0.09336	0.14196	0.49119	0.18287	0.02001	0.14936	0.19893
x_{fil}	0.00009	0.00045	0.00019	0.00021	-0.00047	0.00013	-0.00018
y_{fil}	0.00044	0.00013	0.00012	-0.00008	0.00002	0.00008	0.00003
z_{fil}	0.00017	-0.00023	0.00006	0.00035	0.00013	0.00024	-0.00017
SVMg_{fil}	0.48724	0.54195	0.64065	0.67453	0.60313	0.66682	0.74928
RUF	-0.34309	-0.53341	-0.06617	-0.12075	-0.14713	-0.21601	-0.03225
%HRR	0.79809	0.67843	0.77200	0.91921	0.84761	0.76485	0.77491

The results of the HAR classifiers are shown in Tables IV and V, and Figs. 4 and 5. Table IV shows the mean and the standard deviation of the accuracy obtained in the cross-validation of all the ML models, considering all the activities. Fig. 4 shows the boxplots of these accuracies. The different columns of the Table and the subfigures show the different configurations of features. For example, in Fig. 4a, the average accuracy of the seven classifiers using only raw acceleration (x, y, z; Fig. 4b) is 0.166. Adding the SVMg feature (x, y, z, SVMg; Fig. 4b) raises the mean accuracy to 0.199. Replacing raw acceleration with its filtered components and magnitude (x_{fil}, y_{fil}, z_{fil}, SVMg_{fil}; Fig. 4c) further increases the mean to 0.244. Using only SVMg_{fil} and the rate-of-change feature RUF (SVMg_{fil}, RUF; Fig. 4d) yields a mean of 0.256, while including %HRR alongside SVMg_{fil} and RUF (SVMg_{fil}, RUF, %HRR; Fig. 4e) produces a mean accuracy of 0.351. However, combining all features (all; Fig. 4f) reduces the mean to 0.306. These results indicate that adding %HRR to SVMg_{fil} and RUF achieves the greatest performance improvement, whereas the full feature set causes a slight drop in accuracy. We have not performed feature selection because it may miss interactions between features. Instead, we performed an ablation test, which better provides insights into the importance of feature groups and interactions between features. Regarding the different ML algorithms, although this is not the aim of this study, we can observe that, in general, LR, LDA, and NB perform better than KNN and DT-CART in most of the configurations. For example, Table IV shows that LR improves the accuracy by 93.33% with features SVMg_{fil} and RUF with respect to x, y, z. When adding the %HRR the improvement with respect to x, y, z is of 166.67%. Moreover, the DL algorithms, TabNet and TabTransformer, yielded inferior results compared to these classical models. This is likely due to the limited dataset size, which restricts the ability of such complex models to generalize effectively.

We have also performed the same experiments but only discriminating between a set of four activities that have been difficult to discern in the literature, such as household chores, climbing and descending the stairs, and an activity that is not difficult to distinguish as a base, in our case, regular walk. In particular, we choose the activities of drawing, dishes, regular pace walking, and climbing/descending stairs. The results are shown in Table V and Fig. 5.

Regarding the individual activities and how they improve with the new features, we have plotted in Fig. 6 the confusion matrix for one of the models, LR; the rest of the models have similar behavior. On the left-hand side of the figure, we show the matrix for the experiment with all activities, and on the right-hand side, the matrix for the experiment with only the 4 activities. We have selected 3 groups of features: at the top of the figure, the raw data; in the middle, SVMg_{fil} and RUF; and at the bottom, SVMg_{fil}, RUF, and %HRR. We can see that, in general, accuracy improves when we add new features, as

indicated by the increase in diagonal values of the matrix. In particular, for all activities, we observe that the activities showing the greatest improvements are washing dishes, which, with raw data, has very poor results, as well as ascending and descending stairs. With only 4 activities, the greatest improvement is seen in the stair activities, whereas %HRR is particularly important for walking.

V. DISCUSSION

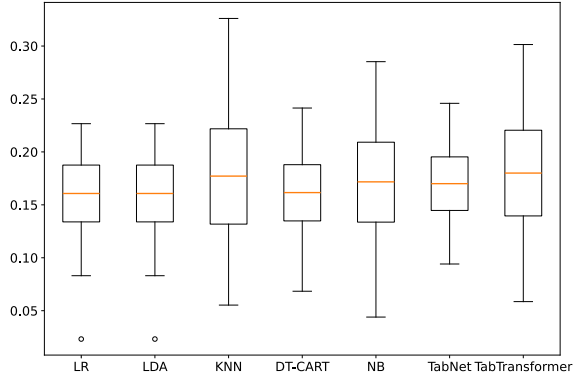
In the correlation matrix shown in Fig. 3, the filtered SVMg appears to play an important role in detecting the activity. We can also see that the heart rate and the HRR also play an important role, with a correlation above 77%. It is worth noting that even though the components y and z of the acceleration correlate, the non-filtered SVMg has a low correlation. This correlation depends highly on the subject, as it is shown in Table III, although all the subjects seem to present a correlation with the y component of the acceleration, most of them (but not all of them) show a correlation with the z component and only a few of them show a correlation with the x component. We can appreciate that in all cases, the correlation with the non-filtered SVMg is very low, contrary to the correlation with the filtered SVMg_{fil}, which is high. This contrasts with the low correlation between the activities and the filtered components of the acceleration (x_{fil}, y_{fil}, z_{fil}), which is in accordance with previous studies, in which they state that filtering the acceleration degrades the results when using each component directly without using SVMg [24].

As we can see, the results of the raw data x,y,z are not good, but if we add only one feature, which is the SVMg to the raw data, as in Fig. 4b, we can see an improvement in the accuracy in all the models. The improvement could be because the SVMg reduces the impact of sensor orientation on the activity discrimination [26]. If instead of using the raw acceleration, we use the filtered acceleration in each component and the SVMg_{fil}, as in Fig. 4c, we see an improvement in the accuracy. The filtered acceleration is a better feature than the raw acceleration. Only with the features of SVMg_{fil} and the RUF can we appreciate a bigger improvement in Fig. 4d. Hence, the ratio between the magnitudes of the unfiltered and filtered acceleration plays an important role in the HAR. If we add to these two features the %HRR, as in Fig. 4e, we can appreciate a bigger improvement in all the models. The heart rate and, by extension, the %HRR are important features in HAR; however, not all devices provide the heart rate continuously. With all the features, such as in Fig. 4f, the accuracy degrades again. This is due to the noise added by some features. If future studies want to use all the features, they must use an algorithm for feature selection before applying the classifier.

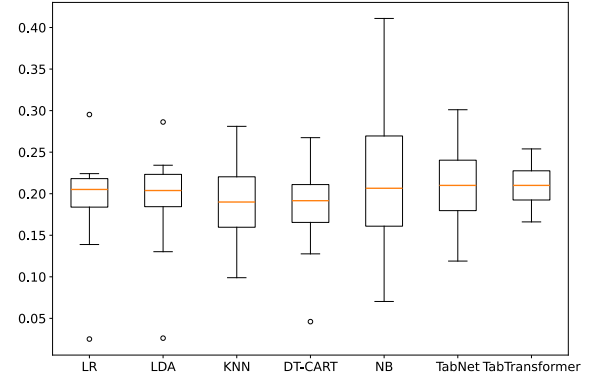
We can see that the discussion for Fig. 4 also applies to Fig. 5, but with better accuracies, as Fig. 5 uses fewer activities; only four activities

TABLE IV. ACCURACY OF THE DIFFERENT ML MODELS, WITH DIFFERENT FEATURES, TAKING INTO ACCOUNT ALL THE ACTIVITIES: LYING IN BED; SITTING IN A CHAIR; DRAWING; WASHING DISHES; SWEEPING THE FLOOR; DESCENDING STAIRS; WALKING SLOWLY; WALKING AT A REGULAR PACE; BRISK WALKING; CLIMBING STAIRS

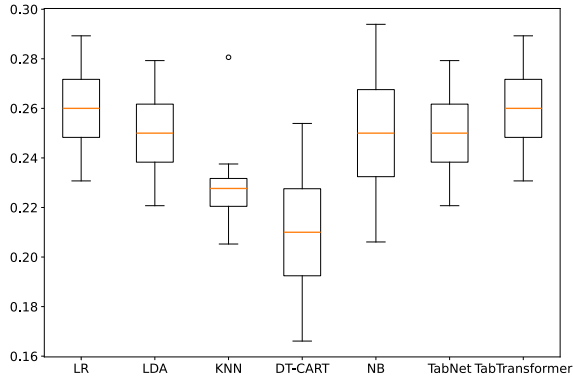
Model	Features					
	x, y, z	x,y,z, SVMg	xfil, yfil, zfil, SVMg _{fil}	SVMg _{fil} RUF	SVM _{gfil} RUF,%HRR	all
LR	0.15 ± 0.06	0.19 ± 0.07	0.26 ± 0.02	0.29 ± 0.04	0.40 ± 0.11	0.31 ± 0.05
LDA	0.15 ± 0.06	0.19 ± 0.07	0.25 ± 0.02	0.26 ± 0.04	0.40 ± 0.10	0.32 ± 0.08
KNN	0.18 ± 0.08	0.19 ± 0.06	0.23 ± 0.02	0.26 ± 0.02	0.31 ± 0.09	0.30 ± 0.09
DT-CART	0.16 ± 0.05	0.18 ± 0.06	0.21 ± 0.03	0.24 ± 0.03	0.30 ± 0.09	0.26 ± 0.08
NB	0.17 ± 0.07	0.22 ± 0.10	0.25 ± 0.03	0.26 ± 0.05	0.35 ± 0.08	0.36 ± 0.08
TabNet	0.17 ± 0.05	0.21 ± 0.06	0.25 ± 0.02	0.26 ± 0.04	0.39 ± 0.09	0.33 ± 0.06
TabTransformer	0.18 ± 0.08	0.21 ± 0.03	0.26 ± 0.02	0.22 ± 0.02	0.31 ± 0.10	0.26 ± 0.09



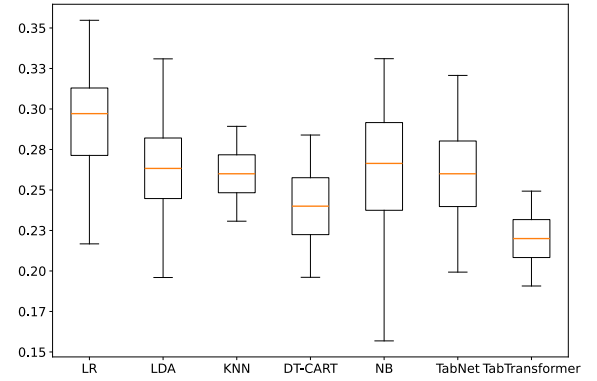
(a) With only the features of the raw components of the acceleration: x, y and z



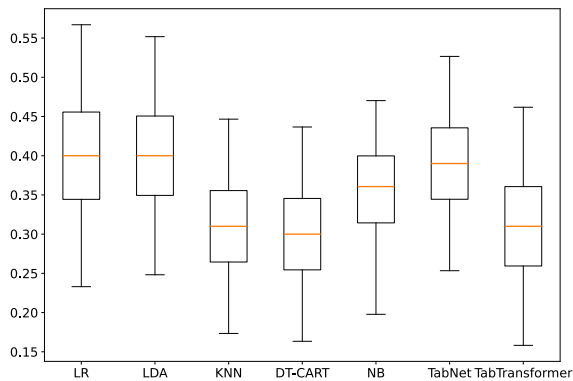
(b) With only the raw acceleration features and sum vector magnitude: x, y, z and SVMg



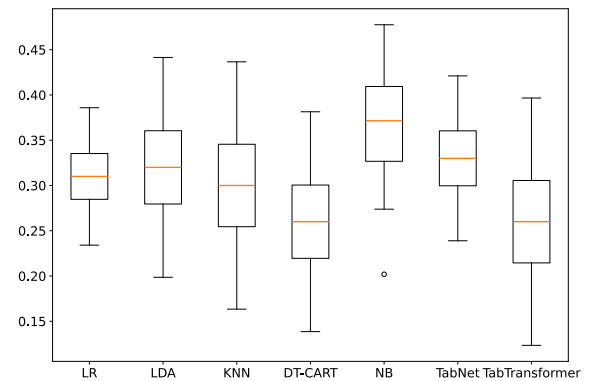
(c) With the filtered acceleration features and its sum vector magnitude: xfil, yfil, zfil and SVM_{gfil}



(d) With the features related to the filtered acceleration and ratio unfiltered-filtered: SVM_{gfil} and RUF



(e) With the features normally used to calculate the MET: SVM_{gfil} RUF and %HRR

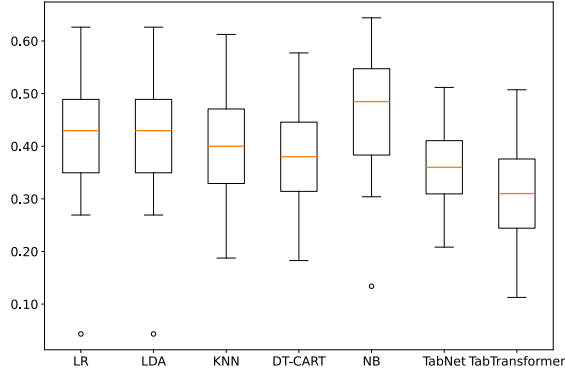


(f) With all the features: x, y, z, SVMg, xfil, yfil, zfil, SVM_{gfil} RUF, HR, and %HRR

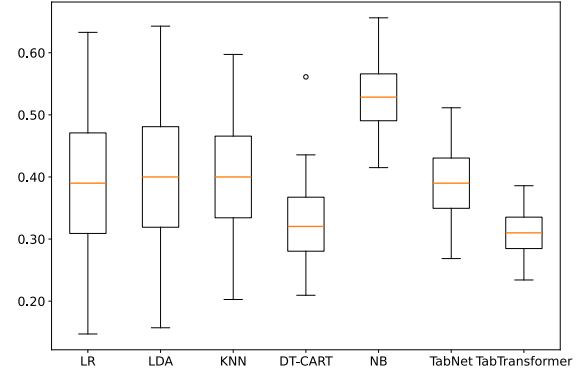
Fig. 4. Accuracy of the different ML models, with different features, taking into account all the activities: lying in bed; sitting in a chair; drawing; washing dishes; sweeping the floor; descending stairs; walking slowly; walking at a regular pace; brisk walking; climbing stairs.

TABLE V. ACCURACY OF THE DIFFERENT ML MODELS, WITH DIFFERENT FEATURES, TAKING INTO ACCOUNT ONLY SOME ACTIVITIES DIFFICULT TO DISCERN: DRAW, DISHES, REGULAR PACE WALKING AND UP-DOWN STAIRS

Model	Features					
	x, y, z	x,y,z, SVMg	xfil, yfil, zfil, SVMg _{fil}	SVMg _{fil} RUF	SVMg _{fil} RUF,%HRR	all
LR	0.40 ± 0.16	0.39 ± 0.16	0.48 ± 0.03	0.51 ± 0.05	0.70 ± 0.11	0.65 ± 0.11
LDA	0.40 ± 0.16	0.40 ± 0.16	0.45 ± 0.02	0.48 ± 0.07	0.70 ± 0.15	0.71 ± 0.11
KNN	0.40 ± 0.14	0.40 ± 0.13	0.50 ± 0.02	0.48 ± 0.06	0.56 ± 0.18	0.58 ± 0.13
DT-CART	0.38 ± 0.13	0.34 ± 0.10	0.46 ± 0.04	0.43 ± 0.06	0.49 ± 0.23	0.43 ± 0.23
NB	0.45 ± 0.15	0.53 ± 0.07	0.49 ± 0.04	0.49 ± 0.05	0.67 ± 0.13	0.67 ± 0.13
TabNet	0.36 ± 0.10	0.39 ± 0.08	0.46 ± 0.03	0.48 ± 0.05	0.68 ± 0.12	0.62 ± 0.12
TabTransformer	0.31 ± 0.13	0.31 ± 0.05	0.40 ± 0.05	0.37 ± 0.02	0.47 ± 0.12	0.45 ± 0.13



(a) With only the features of the raw components of the acceleration: x, y and z



(b) With only the raw acceleration features and sum vector magnitude: x, y, z and SVMg

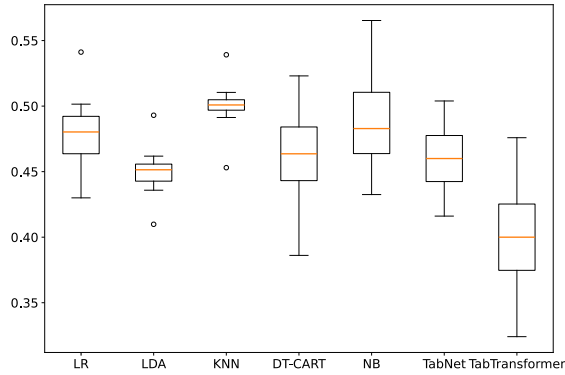
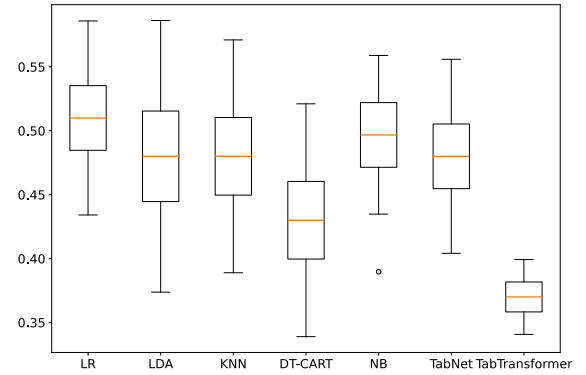
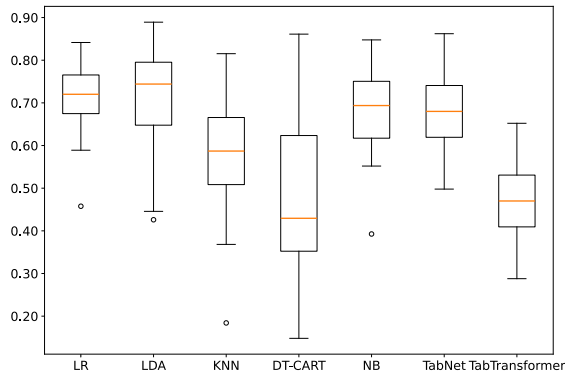
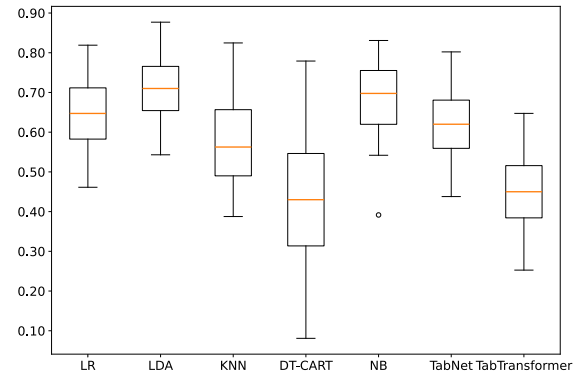

 (c) With the filtered acceleration features and its sum vector magnitude: xfil, yfil, zfil and SVMg_{fil}

 (d) With the features related to the filtered acceleration and ratio unfiltered-filtered: SVMg_{fil} and RUF

 (e) With the features normally used to calculate the MET: SVMg_{fil}, RUF and %HRR

 (f) With all the features: x, y, z, SVMg, xfil, yfil, zfil, SVMg_{fil}, RUF, HR, and %HRR

Fig. 5. Accuracy of the different ML models, with different features, taking into account only some activities difficult to discern: draw, dishes, regular pace walking and up-down stairs.

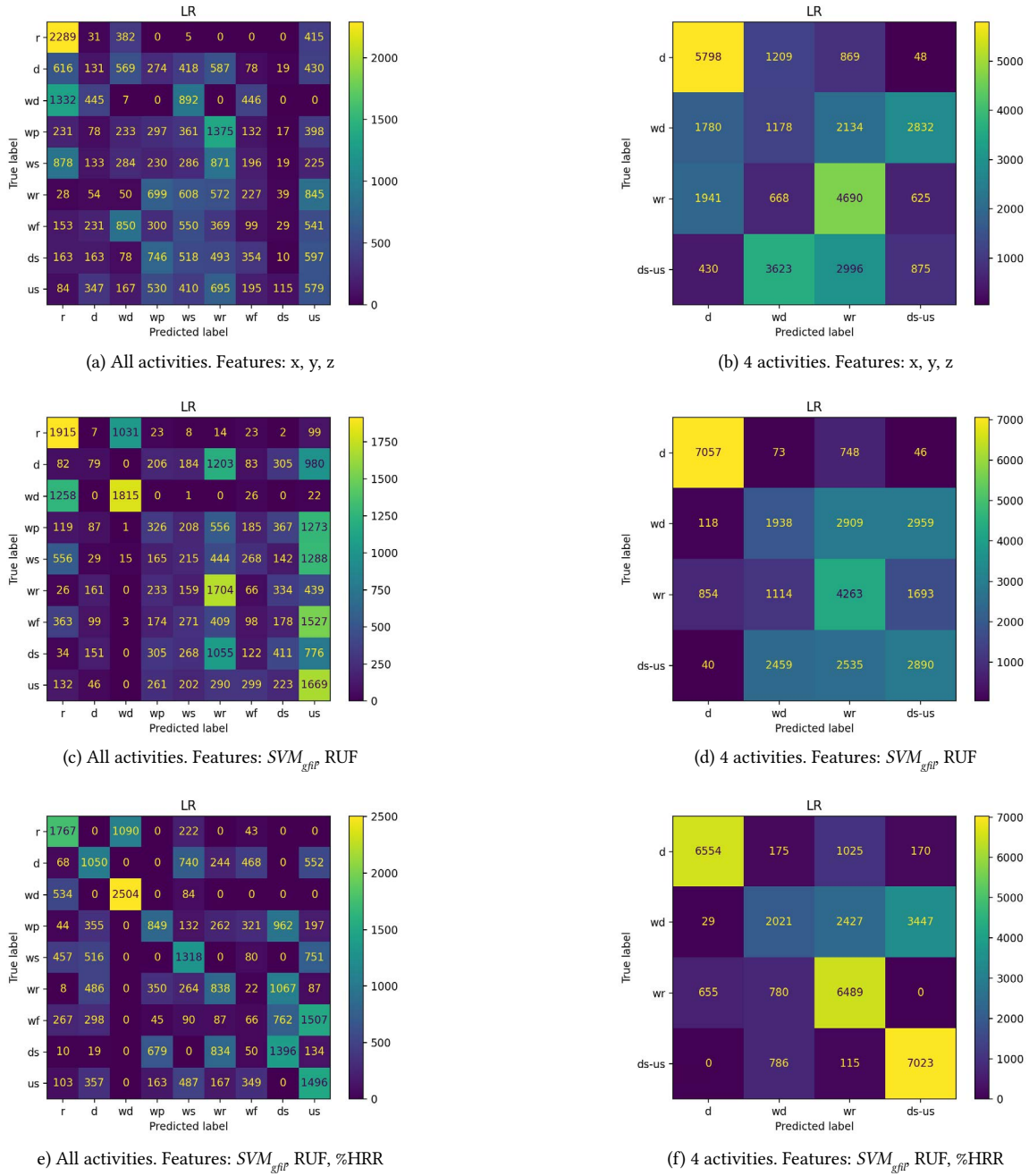


Fig. 6. Confusion matrix of one of the ML models, LR, shows the performance in each activity individually. The left-hand side shows the experiment with all the activities, and the right-hand side shows the experiments with 4 activities: drawing, dishes, regular pace walking, and up-down stairs. The top figures show the experiments with the raw data: x, y, z. The middle figures show the experiments with the features SVM_{gfil} and RUF, and the bottom figures show the experiments with features SVM_{gfil} , RUF, and %HRR.

compared to the complete set of ten activities. In Table IV, we can observe that, for example, LR shows a 27.5% accuracy improvement with features SVM_{gfil} and RUF compared to x, y, and z. When adding the %HRR, the improvement compared to x, y, and z is 75%. As we can see in both figures, the use of features related to MET, such as SVM_{gfil} , RUF, and %HRR can considerably improve the accuracy of the HAR. For example, with the models LR and LDA in Fig. 5, we can appreciate that using these three features almost doubles the accuracy (from 40% to 70%) of using only the raw acceleration components.

The findings of this study highlight the significant impact of incorporating novel features such as the RUF, %HRR, and filtered acceleration on the accuracy of HAR models. These features contribute to a more robust representation of physiological and movement

data, enhancing the model's ability to distinguish between different activities. The generalizability of our methodology is promising, as these features are derived from fundamental physiological and kinematic principles that are not specific to a particular dataset or scenario. By normalizing individual variations and focusing on relative changes in activity intensity, our approach can be adapted to various contexts beyond the specific dataset used in this study. For instance, in different HAR datasets with varying participant demographics, sensor placements, and activity types, the inclusion of RUF, %HRR, and filtered acceleration can similarly improve model performance. Hence, by applying the proposed feature set to diverse datasets, we can extend our findings' applicability to broader real-world settings, such as clinical monitoring, fitness tracking, and eldercare.

A. Limitations and Future Works

The validation method, LOSO-CV, helps evaluate the model's ability to generalize across individuals, considering inter-subject variability. However, it poses the challenge of generalizing across subjects with very different activity patterns. The model may perform inconsistently if there is high variability, indicating potential overfitting to specific individuals' data. This challenge is mitigated by using %HRR instead of HR and the RUF, which normalize the data between individuals. Furthermore, LOSO-CV is computationally expensive compared to 5-fold validation, especially if the number of subjects is high. This is because it requires training and evaluating one model per subject (tested with one subject and trained with the rest).

Although we have not used any windowing or any feature extraction on each feature, such as the standard temporal and frequency features: mean, average, sleekness, entropy, etc., our results could help other studies to windowing and extract the standard features over $SVMg_{fil}$, RUF and %HRR, and improve the HAR detection. Using such feature extraction would make our results less transparent and more dependent on the feature extraction than on the feature manual selection, hiding our proposal.

VI. CONCLUSION

This study demonstrates that incorporating features related to the MET, such as the $SVMg_{fil}$, the RUF, and the %HRR, significantly improves the accuracy of HAR models. These features provide a more robust representation of movement and physiological data, allowing for better differentiation of challenging activities, such as stair climbing and household chores, which are difficult to distinguish using only raw acceleration data.

Our results show that using these features leads to accuracy improvements of up to 166.67%, especially for stair climbing and household chores. This highlights the importance of feature selection and preprocessing in HAR tasks. The integration of heart rate data with acceleration data enhances the detection of activity intensity, demonstrating the value of combining physiological metrics with movement data for more precise activity recognition.

By leveraging these features, our approach contributes to developing more reliable HAR systems with potential applications in healthcare, fitness tracking, and everyday activity monitoring.

CREDiT AUTHORSHIP CONTRIBUTION STATEMENT

Maria Bermudez-Edo:Methodology, Validation, Investigation, Resources, Software, Writing- Original Draft, Supervision, Project Administration.

Alberto Durán López:Investigation, Visualization, Writing-Review & Editing.

Daniel Bolaños-Martínez:Investigation, Visualization, Writing-Review & Editing.

Jose Luis Garrido:Investigation, Writing- Review & Editing.

Maria Jose Rodriguez-Fortiz:Investigation, Writing- Review & Editing, Project Administration.

DATA STATEMENT

All data used in this paper can be found in <https://doi.org/10.5281/zenodo.17671225>

DECLARATION OF CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

COMPLIANCE WITH ETHICAL STANDARDS

All participants provided informed consent prior to data collection, and confidentiality was maintained throughout the study by anonymizing all personal data. The research protocol was approved by the Ethics Committee of the University of Granada (approval number 4400/CEIH/2024) and was conducted according to the principles of the Declaration of Helsinki.

ACKNOWLEDGMENT

Grant PID2023-149185OB-I00 project funded by MICIU/AEI/10.13039/501100011033 and co-funded by ERDF/EU; project C-SEJ-128-UGR23 funded by Consejería de Universidad, Investigación e Innovación and by ERDF Andalusia Program 2021–2027.

REFERENCES

- [1] J. F. Sallis, D. Adlakha, A. Oyeyemi, D. Salvo, "An international physical activity and public health research agenda to inform coronavirus disease-2019 policies and practices," *Journal of sport and health science*, vol. 9, no. 4, pp. 328–334, 2020, doi: <https://doi.org/10.1016/j.jshs.2020.05.005>
- [2] M. Izquierdo, R. Merchant, J. Morley, S. Anker, I. Aprahamian, H. Arai, M. Aubertin-Leheudre, R. Bernabei, E. Cadore, M. Cesari, others, "International exercise recommendations in older adults (ICFSR): expert consensus guidelines," *The journal of nutrition, health & aging*, vol. 25, no. 7, pp. 824–853, 2021, doi: <https://doi.org/10.1007/s12603-021-1665-8>
- [3] K. Ito, Y. Suzuki, T. Sakaguchi, K. Fujiwara, Y. Nishii, H. Yasui, O. Taguchi, O. Hataji, "Physical Activity Using a Wearable Device as an Alternative to Performance Status in Patients With Advanced Lung Cancer," *JAMA oncology*, 2024, doi: <https://doi.org/10.1001/jamaoncol.2024.0023>
- [4] N. Aguilar-Farias, W. J. Brown, T. L. Skinner, G. G. Peeters, "Metabolic equivalent values of common daily activities in middle-age and older adults in free-living environments: a pilot study," *Journal of Physical Activity and Health*, vol. 16, no. 3, pp. 222–9, 2019, doi: <https://doi.org/10.1123/jpah.2016-0400>
- [5] A. Davoudi, M. T. Mardini, D. Nelson, F. Albinali, S. Ranka, P. Rashidi, T. M. Manini, others, "The effect of sensor placement and number on physical activity recognition and energy expenditure estimation in older adults: Validation study," *JMIR mHealth and uHealth*, vol. 9, no. 5, p. e23681, 2021, doi: <https://doi.org/10.2196/23681>
- [6] F. Tokmak, T. Koivisto, O. Lahdenoja, T. Vasankari, S. Jaakkola, K. E. J. Airaksinen, "Mechanocardiography detects improvement of systolic function caused by resynchronization pacing," *Physiological Measurement*, vol. 44, no. 12, p. 125009, 2023, doi: <https://doi.org/10.1088/1361-6579/ad1197>
- [7] M. Sevil, M. Rashid, Z. Maloney, I. Hajizadeh, S. Samadi, M. R. Askari, N. Hobbs, R. Brandt, M. Park, L. Quinn, others, "Determining physical activity characteristics from wristband data for use in automated insulin delivery systems," *IEEE sensors journal*, vol. 20, no. 21, pp. 12859–12870, 2020, doi: <https://doi.org/10.1109/JSEN.2020.3000772>
- [8] F. Tokmak, B. Semiz, "Unveiling the Relationships Between Seismocardiogram Signals, Physical Activity Types and Metabolic Equivalent of Task Scores," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 2, pp. 479–487, 2022, doi: <https://doi.org/10.1109/TBME.2022.3194594>
- [9] V. Dentamaro, V. Gattulli, D. Impedovo, F. Manca, "Human activity recognition with smartphone-integrated sensors: A survey," *Expert Systems with Applications*, p. 123143, 2024, doi: <https://doi.org/10.1016/j.eswa.2024.123143>
- [10] A. Uphill, K. L. Kendall, A. Fogarty, S. Guppy, H. Brown, T. Zomer, S. Parker, G. G. Haff, "Validity of Apple Watch, Garmin Forerunner 935 and GENEActiv for estimating energy expenditure during close quarter battle training in Special Forces soldiers," *European Journal of Sport Science*, 2024, doi: <https://doi.org/10.1002/ejsc.12080>

- [11] M. L. Danielsson, M. Vergeer, G. Plasqui, J. K. Baumgart, "Accuracy of the Apple Watch Series 4 and Fitbit Versa for Assessing Energy Expenditure and Heart Rate of Wheelchair Users During Treadmill Wheelchair Propulsion: Cross-sectional Study," *JMIR Formative Research*, vol. 8, no. 1, p. e52312, 2024, doi: <https://doi.org/10.2196/52312>
- [12] M. Jetté, K. Sidney, G. Blümchen, "Metabolic equivalents (METs) in exercise testing, exercise prescription, and evaluation of functional capacity," *Clinical cardiology*, vol. 13, no. 8, pp. 555–565, 1990, doi: <https://doi.org/10.1002/clc.4960130809>
- [13] U. Verma, P. Tyagi, M. Kaur, "Artificial intelligence in human activity recognition: a review," *International Journal of Sensor Networks*, vol. 41, no. 1, pp. 1–22, 2023, doi: <https://doi.org/10.1504/IJSNET.2023.128503>
- [14] M. Nakanishi, S. Izumi, S. Nagayoshi, H. Kawaguchi, M. Yoshimoto, T. Shiga, T. Ando, S. Nakae, C. Usui, T. Aoyama, others, "Estimating metabolic equivalents for activities in daily life using acceleration and heart rate in wearable devices," *Biomedical engineering online*, vol. 17, no. 1, pp. 1–18, 2018, doi: <https://doi.org/10.1186/s12938-018-0532-2>
- [15] C. Lütznier, H. Voigt, I. Roeder, S. Kirschner, J. Lütznier, "Placement makes a difference: accuracy of an accelerometer in measuring step number and stair climbing," *Gait & posture*, vol. 39, no. 4, pp. 1126–1132, 2014, doi: <https://doi.org/10.1016/j.gaitpost.2014.01.022>
- [16] D. R. Witt, R. A. Kellogg, M. P. Snyder, J. Dunn, "Windows into human health through wearables data analytics," *Current Opinion in Biomedical Engineering*, vol. 9, pp. 28–46, 2019, doi: <https://doi.org/10.1016/j.cobme.2019.01.001>
- [17] N. M. Remillard, W. Hoogkamer, G. Layec, J. Sirard, "Validation of Exercise and Postexercise Energy Expenditure Estimates Using Hip and Wrist Accelerometer Algorithms," *Journal for the Measurement of Physical Behaviour*, vol. 7, no. 1, 2024, doi: <https://doi.org/10.1123/jmpb.2024-0015>
- [18] J. Xie, D. Wen, L. Liang, Y. Jia, L. Gao, J. Lei, others, "Evaluating the validity of current mainstream wearable devices in fitness tracking under various physical activities: comparative study," *JMIR mHealth and uHealth*, vol. 6, no. 4, p. e9754, 2018, doi: <https://doi.org/10.2196/mhealth.9754>
- [19] A. Shcherbina, C. M. Mattsson, D. Waggott, H. Salisbury, J. W. Christle, T. Hastie, M. T. Wheeler, E. A. Ashley, "Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort," *Journal of personalized medicine*, vol. 7, no. 2, p. 3, 2017, doi: <https://doi.org/10.3390/jpm7020003>
- [20] K. Ohkawara, Y. Oshima, Y. Hikiyama, K. Ishikawa-Takata, I. Tabata, S. Tanaka, "Real-time estimation of daily physical activity intensity by a triaxial accelerometer and a gravity-removal classification algorithm," *British Journal of Nutrition*, vol. 105, no. 11, pp. 1681–1691, 2011, doi: <https://doi.org/10.1017/S0007114510005441>
- [21] N. Koizumi, H. Ogata, Y. Negishi, H. Nagayama, M. Kaneko, K. Kiyono, N. Omi, "Energy Expenditure of Disaster Relief Operations Estimated Using a Tri-Axial Accelerometer and a Wearable Heart Rate Monitor," *International Journal of Environmental Research and Public Health*, vol. 20, no. 9, p. 5742, 2023, doi: <https://doi.org/10.3390/ijerph20095742>
- [22] T. Chen, S. Chen, T. Honda, H. Kishimoto, Y. Nofuji, K. Narazaki, "Accelerometer-measured sedentary behavior and risk of functional disability in older Japanese adults: a 9-year prospective cohort study," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 20, no. 1, p. 91, 2023, doi: <https://doi.org/10.1186/s12966-023-01490-6>
- [23] I. M. Pires, F. Hussain, G. Marques, N. M. Garcia, "Comparison of machine learning techniques for the identification of human activities from inertial sensors available in a mobile device after the application of data imputation techniques," *Computers in Biology and Medicine*, vol. 135, p. 104638, 2021, doi: <https://doi.org/10.1016/j.combiomed.2021.104638>
- [24] Ç. B. Erda, I. Atasoy, K. Açıci, H. O., "Integrating features for accelerometer-based activity recognition," *Procedia Computer Science*, vol. 98, pp. 522–527, 2016, doi: <https://doi.org/10.1016/j.procs.2016.09.070>
- [25] A. K. Panja, A. Rayala, A. Agarwala, S. Neogy, C. Chowdhury, "A hybrid tuple selection pipeline for smartphone based Human Activity Recognition," *Expert Systems with Applications*, vol. 217, p. 119536, 2023, doi: <https://doi.org/10.1016/j.eswa.2023.119536>
- [26] A. Poli, G. Cosoli, L. Scalise, S. Spinsante, "Impact of wearable measurement properties and data quality on ADLs classification accuracy," *IEEE Sensors Journal*, vol. 21, no. 13, pp. 14221–14231, 2020, doi: <https://doi.org/10.1109/JSEN.2020.3009368>
- [27] O. Barut, L. Zhou, Y. Luo, "Multitask LSTM model for human activity recognition and intensity estimation using wearable sensor data," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8760–8768, 2020, doi: <https://doi.org/10.1109/JIOT.2020.2996578>
- [28] T. Stewart, A. Narayanan, L. Hedayatrad, J. Neville, L. Mackay, S. Duncan, "A Dual-Accelerometer System for Classifying Physical Activity in Children and Adults," *Medicine and science in sports and exercise*, vol. 50, no. 12, pp. 2595–2602, 2018, doi: <https://doi.org/10.1249/mss.0000000000001717>
- [29] J. Lu, X. Zheng, M. Sheng, J. Jin, S. Yu, "Efficient human activity recognition using a single wearable sensor," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11137–11146, 2020, doi: <https://doi.org/10.1109/JIOT.2020.2995940>
- [30] F. M. Garcia-Moreno, M. J. Rodriguez-Fortiz, P. Barnaghi, M. Bermudez-Edo, "Modelling time-series data generation with diffusion models for triaxial data," *Applied Soft Computing*, p. 114195, 2025, doi: <https://doi.org/10.1016/j.asoc.2025.114195>
- [31] E. Ferrara, "Large language models for wearable sensor-based human activity recognition, health monitoring, and behavioral modeling: A survey of early trends, datasets, and challenges," *Sensors*, vol. 24, no. 15, p. 5045, 2024, doi: <https://doi.org/10.3390/s24155045>
- [32] N. Gupta, S. K. Gupta, R. K. Pathak, V. Jain, P. Rashidi, J. S. Suri, "Human activity recognition in artificial intelligence framework: a narrative review," *Artificial intelligence review*, vol. 55, no. 6, pp. 4755–4808, 2022, doi: <https://doi.org/10.1007/s10462-021-10116-x>
- [33] P. Kumar, S. Chauhan, L. K. Awasthi, "Human activity recognition (har) using deep learning: Review, methodologies, progress and future research directions," *Archives of Computational Methods in Engineering*, vol. 31, no. 1, pp. 179–219, 2024, doi: <https://doi.org/10.1007/s11831-023-09986-x>
- [34] S. Zhang, Y. Li, S. Zhang, F. Shahabi, S. Xia, Y. Deng, N. Alshurafa, "Deep learning in human activity recognition with wearable sensors: A review on advances," *Sensors*, vol. 22, no. 4, p. 1476, 2022, doi: <https://doi.org/10.3390/s22041476>
- [35] M. Liandana, D. P. Hostiadi, G. A. Pradipta, "A New Approach for Human Activity Recognition (HAR) Using A Single Triaxial Accelerometer Based on a Combination of Three Feature Subsets," *International Journal of Intelligent Engineering & Systems*, vol. 17, no. 2, 2024, doi: <https://doi.org/10.22266/ijies2024.0430.21>
- [36] N. M. Obiri, K. Van Laerhoven, "A Survey of LoRaWAN-integrated Wearable Sensor Networks for Human Activity Recognition: Applications, Challenges and Possible Solutions," *IEEE Open Journal of the Communications Society*, 2024, doi: <https://doi.org/10.1109/OJCOMS.2024.3484002>
- [37] Y. Li, G. Yang, Z. Su, S. Li, Y. Wang, "Human activity recognition based on multi-environment sensor data," *Information Fusion*, vol. 91, pp. 47–63, 2023, doi: <https://doi.org/10.1016/j.inffus.2022.10.015>
- [38] L. Liu, Y. Sun, X. Ge, "A Hybrid Multi-Person Fall Detection Scheme Based on Optimized YOLO and ST-GCN," *International Journal of Interactive Multimedia & Artificial Intelligence*, vol. 9, no. 2, 2025, doi: <https://doi.org/10.9781/ijimai.2024.09.003>
- [39] C. Contoli, V. Freschi, E. Lattanzi, "Energy-aware human activity recognition for wearable devices: A comprehensive review," *Pervasive and Mobile Computing*, vol. 104, p. 101976, 2024, doi: <https://doi.org/10.1016/j.pmcj.2024.101976>
- [40] M. Heydarian, T. E. Doyle, "rWISDM: Repaired WISDM, a Public Dataset for Human Activity Recognition," *arXiv preprint arXiv:2305.10222*, 2023, doi: <https://doi.org/10.48550/arXiv.2305.10222>
- [41] M. Arshad, F. H. Jaskani, M. A. Sabri, F. Ashraf, M. Farhan, M. Sadiq, H. Raza, "Hybrid machine learning techniques to detect real time human activity using UCI dataset," *EAI Endorsed Transactions on Internet of Things*, vol. 7, no. 26, pp. e1–e1, 2021, doi: <https://doi.org/10.4108/eai.26-5-2021.170006>
- [42] D. Micucci, M. Mobilio, P. Napolitano, "Unimib shar: A dataset for human activity recognition using acceleration data from smartphones," *Applied Sciences*, vol. 7, no. 10, p. 1101, 2017, doi: <https://doi.org/10.20944/preprints201706.0033.v2>
- [43] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Padiaditis, M. Tsiknakis, "The mobiaact dataset: Recognition of activities of daily living using smartphones," vol. 2, pp. 143–151, 2016, doi: <https://doi.org/10.5220/0005792401430151>
- [44] N. Sikder, A.-A. Nahid, "KU-HAR: An open dataset for heterogeneous human activity recognition," *Pattern Recognition Letters*, vol. 146, pp. 46–54, 2021, doi: <https://doi.org/10.1016/j.patrec.2021.02.024>

- [45] S. Gashi, C. Min, A. Montanari, S. Santini, F. Kawsar, "A multidevice and multimodal dataset for human energy expenditure estimation using wearable devices," *Scientific Data*, vol. 9, no. 1, p. 537, 2022, doi: <https://doi.org/10.1038/s41597-022-01643-5>
- [46] S. Chan, Y. Hang, C. Tong, A. Acquah, A. Schonfeldt, J. Gershuny, A. Doherty, "CAPTURE-24: A large dataset of wrist-worn activity tracker data collected in the wild for human activity recognition," *Scientific Data*, vol. 11, no. 1, p. 1135, 2024, doi: <https://doi.org/10.1038/s41597-024-03960-3>
- [47] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE transactions on information technology in biomedicine*, vol. 10, no. 1, pp. 156–167, 2006, doi: <https://doi.org/10.1109/TITB.2005.856864>
- [48] S. Ö. Arik, T. Pfister, "Tabnet: Attentive interpretable tabular learning," vol. 35, no. 8, pp. 6679–6687, 2021, doi: <https://doi.org/10.1609/aaai.v35i8.16826>
- [49] X. Huang, A. Khetan, M. Cvitkovic, Z. Karnin, "Tabtransformer: Tabular data modeling using contextual embeddings," *arXiv preprint arXiv:2012.06678*, 2020, doi: <https://doi.org/10.48550/arXiv.2012.06678>



Maria Jose Rodriguez-Fortiz

Maria Jose Rodriguez-Fortiz} {Maria Jose received the B.S. and Ph.D. degrees in Computer Science from the University of Granada, Spain. Since 1990, she has been a Lecturer with the Department of Languages and Computer Systems, University of Granada. She has served as a lead researcher and has participated in several research projects at national, regional, and local levels, as well as in industrial contracts. Her research focuses on the application of ICT to individuals with special needs and older adults, specializing in software accessibility, usability, and adaptation based on user profiles and contextual information.



Maria Bermudez-Edo

Maria received the Ph.D. degree in Information and Communication Technologies from the University of Granada, Spain, in 2013. She was a Research Fellow with the University of Surrey, Guildford, U.K., with two previous secondments to the University of California at Berkeley, CA, USA, and ETH Zurich, Switzerland. She is currently an Associate Professor with the University of Granada. She was previously with the European Patent Office, The Hague, Netherlands; Siemens, Herentals, Belgium; and Telefónica, Madrid, Spain. Her research interests include semantics and data analytics in IoT, smart cities, and eHealth.



Alberto Durán-López

Alberto received the double B.S. degree in Computer Science and Mathematics from the University of Granada, Spain, in 2020. He holds two master's degrees, in education and mathematics, respectively. His professional career includes roles as a full-stack developer and a secondary school computer science teacher. Currently, he is an Interim Professor with the Department of Software Engineering, University of Granada, where he conducts research in data science for his ongoing Ph.D. thesis. His research interests include data analytics and statistics in IoT and smart cities.



Daniel Bolaños-Martínez

Daniel received the double B.S. degree in Computer Science and Mathematics, and the Ph.D. degree in Computer Science from the University of Granada, Spain, completing the latter in 2025. He has worked as a research assistant on projects focused on data analysis and sensor data collection systems. Since 2023, he has been an Interim Professor with the Department of Software Engineering, School of Computer and Telecommunication Engineering, University of Granada. His research interests include machine learning and statistics in IoT and smart cities.



Jose Luis Garrido

Jose Luis received the B.S. and Ph.D. degrees in Computer Science from the University of Granada, Spain, completing the latter in 2003. He has been a faculty member with the Department of Software Engineering, School of Computer and Telecommunication Engineering, University of Granada, since 1990, where he currently holds the position of Full Professor. His R&D activities focus on the applications of distributed and ubiquitous computing to smart environments and eHealth, with specific interests in pervasive and context-aware computing, collaborative systems, and human-computer interaction.