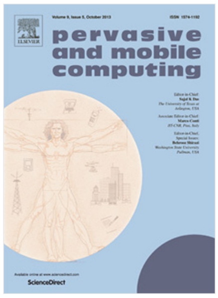
Accepted Manuscript



Design and assessment of the data analysis process for a wrist-worn smart object to detect atomic activities in the smart home

Qin Ni, Ian Cleland, Chris Nugent, Ana Belén García Hernando, Iván Pau de la Cruz

PII: S1574-1192(17)30576-X
DOI: https://doi.org/10.1016/j.pmcj.2019.03.006
Reference: PMCJ 1009

To appear in: *Pervasive and Mobile Computing*

Received date : 22 November 2017
Revised date : 5 January 2019
Accepted date : 30 March 2019

Design and assessment of the Data Analysis process for a wrist-worn smart object to detect atomic activities in the smart home

Qin Ni, Ian Cleland, Chris Nugent, Ana Belén García Hernando and Iván Pau de la Cruz

Abstract— **The ability to accurately identify the different activities of daily living (ADLs) is considered as one of the basis to foster new technological solutions inside the smart home. Current ADL recognition proposals, still however, struggle to accurately and robustly identify the range of different activities that can be performed at home, namely static, dynamic and transient activities, and the high variety of technologies and data analysis possibilities to classify the information gathered by the sensors. In this paper, we describe the methodological approach that we have followed for the processing, analysis and classification of data obtained by a simple and non-intrusive smart object with the objective to detect atomic (i.e. non-divisible) activities inside the smart home. The smart object consists of a wrist-worn 3D accelerometer, which presents as its advantages its customizability and usability. We have performed a set of systematic experiments involving ten people and have followed the steps from data gathering to the comparison of different classification techniques, to find out that it is possible to select a complete succession of data processing steps in order to detect, with high accuracy, a set of atomic activities of daily life with the selected smart object, which performs well with different independent datasets besides ours.**

Keywords — **Smart home, smart object, activity recognition, class imbalance, ensemble classification**

This work was supported in part by Smart Environment Research Group of University of Ulster and the research group on Telematics for Knowledge and Information Society (T>SIC) of Universidad Politécnica de Madrid.

Q. Ni is with College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, PR China. (email: niqin@shnu.edu.cn).

I. Cleland and C. Nugent are with the School of Computing and Mathematics, University of Ulster, Belfast, BT37 0QB, U. K. (email: i.cleland@ulster.ac.uk and cd.nugent@ulster.ac.uk).

A. B. García Hernando and I. Pau de la Cruz are with the Department of Telematic and Electronic Engineering of Universidad Politécnica de Madrid (UPM), Madrid, Spain (e-mail: anabelen.garcia@upm.es and ivan.pau@upm.es).

I. INTRODUCTION

1.1. Context of the research

Recent advances in sensing, networking and ambient intelligence technologies have resulted in a rapid emergence of smart environments. Among these, the Smart Home (SH) has gained a lot of attention for its potential in providing enhanced quality of life within the home. If, based on the SH, it is possible to detect and interpret the behaviour and context of the inhabitants at home, we could think about enhancing their quality of life. This scenario could lead to improvements in independence of people with continuing care needs (such as the elderly) by reducing caregivers' time and healthcare costs in general, without losing the safety that continuous and unobtrusive monitoring provides. The ability to correctly and automatically identify human activities and infer an inhabitant's behaviour at home, has the potential to lend itself to a wide range of applications, such as the detection of health emergencies [1], recommendation services for correct execution of complex activities [2], professional advice on routine lifestyle [3], anomaly detection [4] or help in treatment prescription [5].

The most accepted method to model the inhabitant's behaviour is through the detection of the activities performed at home. There are different types of these activities, but the commonality to all of them is that a non-technical person should be able to recognize them as activities when faced with them. Activities of interest within a SH may include Activities of Daily Life (ADLs) relating to self-care and domestic tasks (sleeping, bathing, dressing, etc.) that are routinely performed by the inhabitant, and ambulatory activities relating to either specific motions or postures of the person (e.g. sitting, standing and transitional activities such as stand-to-sit and sit-to-stand).

In addition to the previous characterization of activities [6], we can consider a different classification of activities into atomic activities, which cannot be further decomposed into simpler activities (recognizable by a non-technical person), and composite activities. This classification is more suitable for engineering purposes, enabling the cooperation of Internet of Things (IoT) and Ambient Intelligence (AmI) technologies to obtain a functional, robust, scalable and reliable solution at home (see Figure 1). The IoT technologies, through the concept of smart objects, can interact physically with the inhabitants in a seamless and unobtrusive way to detect atomic activities. The AmI technologies process and relate the atomic activities produced by the IoT technologies and deal with aspects such as concurrency, uncertainty or context awareness to obtain composite activities used to model the inhabitants' behaviour.

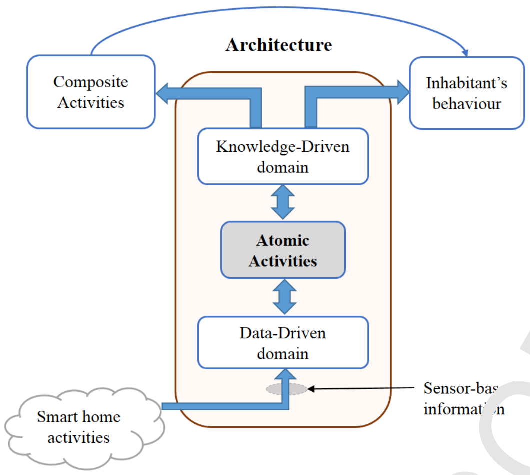


Figure 1. Decoupling of IoT and AmI subsystems through the detection of intermediate atomic activities.

In the proposed approach, illustrated by Figure 1, the atomic activities are the primitives enabling both the modelling of the inhabitant's behaviour (including the composite activity detection) and the improvement of the resilience of the whole system by decoupling both subsystems (IoT and AmI). This paper is focused on the detection of atomic activities through smart objects (IoT subsystem). To improve the recognition of atomic activities, several practical issues must be resolved, this includes issues around usability and acceptability of the solution. In addition, some technical issues difficult to solve for the solution such as the identification of the inhabitant, calibration and power consumption, pose significant challenges.

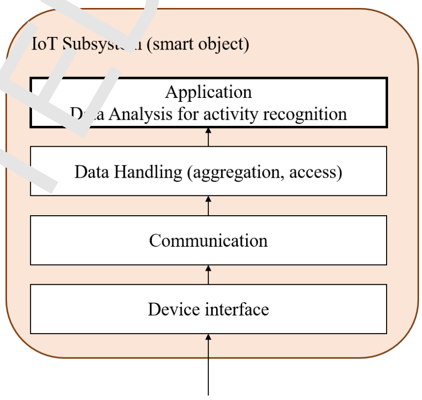


Figure 2. Conceptual model of the IoT subsystem to detect atomic activities.

Figure 2 shows a simplified conceptual model of the IoT subsystem functionalities required for the detection of atomic activities. This subsystem should handle the issues related with the acquisition of raw data from the smart objects, communication among devices (smart objects and other entities), data handling (data hub for the smart home) and the final application (detection of atomic activities). Although these are significant challenges, this paper will focus more on the data analysis procedures for recognizing these "atomic activities" through the use of a specific smart object.

Regarding the detection of the activities, the main issue is the reliable translation from the quantitative data gathered by the sensors to the specific activity being carried out. There are three main issues hindering further improvements in activity recognition performance. Firstly, from the many features that can be calculated, only some will be relevant. Secondly, due to the short duration of transitional activities, there are much less samples of transitional activities (minority classes) than static and dynamic activities (majority classes). Thirdly, a single classifier has limited ability to classify the instances for three classes. Therefore, combining the proper various base classifiers can leverage their strengths and improve the classification performance.

1.2. Objective

This paper describes the methodology we have followed to build the complete Data-driven domain of Figure 1 to accurately detect three types of atomic activities by using a single low-intrusive smart object. The smart object is a wristband provided with a 3D accelerometer. This wristband is a wearable object commonly accepted by most of the population and supports two interesting features: (1), the identification of the inhabitant (each person should have her own wristband). And (2), the customizability of the device (the wristband object provides an opportunity to control the orientation and location of the device on the body, simplifying issues around its calibration).

The atomic activities will be physical activities of the three main types: stationary, dynamic and transitional. The selected activities are considered roughly orthogonal with each other (cannot be performed in parallel with others).

The method to build the data analysis functionality of the wristband will be based on several machine learning techniques and classifiers.

1.3. Summary of the contributions

The main contributions of this paper are:

- To present the technical foundations of the data analysis process from a smart wrist-band that facilitates the recognition of atomic activities in smart home and verify the feasibility of detecting three types of atomic activities by using a single non-intrusive data acquisition device.

- To contribute a benchmark dataset that contains sensor data of static, dynamic as well as transitional activities by using a small, low-cost and non-intrusive wrist-worn accelerometer [7]. The data are collected from a group of participants in a lab setting but includes elements to resemble a real home scenario (they were told what to do but not how to do it, see section 3.1). Specifically, we focus on data collection of transitional activities together with static and dynamic activities, which is often ignored in current research.

- To develop and test a method to compare different machine learning techniques and allow the most adequate machine learning selection for the wrist-worn accelerometer data. By paying attention to the class imbalance in the dataset, which is caused by the very short duration of transitional activities, the framework includes features selection algorithms, oversampling to balance the classes and different classification models, including ensemble approaches. The data processing system selected by following this method has proven to yield good and balanced results also for other researchers' datasets independent from ours.

The rest of the paper is organized as follows: Section II introduces the related work in activity recognition area. Section III describes the experiment protocol and data collection. Sections IV concentrates the methods adopted in the development of the activity recognition approach including signal preprocessing, feature extraction, feature selection, data rebalance, and ensemble classifier construction. Section V discusses the experimental results and the obtained performance on the atomic activities recognition. Finally, section VI provides the conclusions and future work.

II RELATED WORK

Researchers have made significant progress in the area of human activity recognition using wearable accelerometers. Their small size allows accelerometers to be embedded into belts, clothes, glasses, wristwatches, jewellery, shoes and mobile devices to make them easier to wear. For example, Chernbumroong et al. [8] proposed a classification method to monitor and recognize eleven ADLs of an elderly person including feeding, brushing teeth, dressing, washing dishes and so on. They used a single accelerometer embedded into a sport watch, which can be worn by an elderly person in an acceptable and non-intrusive way. Wang et al. [9] used a single accelerometer placed on two different parts of the body: the waist and the left ankle, to monitor eight common domestic activities: sitting, lying, standing up from lying, standing, walking, running, bicycling and jumping. Besides the single accelerometer, multiple accelerometers were placed on distributed body locations to improve the classification accuracy. Gao et al. [10] placed four accelerometers at four different locations of the body (chest, left under-arm, waist and thigh) to monitor and recognize five physical activities. Their system achieved an overall recognition accuracy of 96.4% by adopting the mean and variance features with a Decision Tree classifier. Cleland et al. [11] described an investigation on the optimal placement of accelerometers for the purpose of detecting a range of everyday activities. In their study, six tri-axial accelerometers were worn at distributed body locations including the chest, lower back, hip, thigh, wrist and foot to collect data of seven activities (walking, jogging, sitting, lying, standing, waking upstairs and walking downstairs). The result showed that reasonable activity detection can be achieved using only two accelerometers and that increasing the number of sensors had no significant impact on the accuracy of the classifier. One drawback of using multiple sensors is, however, related to the placement of the accelerometers that would make the activity monitoring highly obtrusive and uncomfortable. A single inertial sensor worn by the user can reduce obtrusiveness to the minimum. In this study, we used one accelerometer placed at the wrist to collect the data. The wrist was chosen as it provides a non-obtrusive location which is comfortable to wear and has shown relatively good accuracy of activity recognition.

A number of activity recognition frameworks with adopting wrist-worn sensors have been proposed. Mehrang et al. [12] presented an assessment of a human activity monitoring framework that covers some of the principal building blocks required for an activity monitoring system, and is comprised of the best preprocessing and parameter choice for an random forest classifier. Qi et al. [13] proposed a two-layer activity recognition framework to classify aerobic, sedentary and free weight activities. However, these frameworks are mainly proposed to monitoring static and dynamic activities, the transitions between multiple activities, which make up a large part of a person's daily activities and can provide additional contextual information for activity recognition, are usually ignored in most current literature due to their short duration. Furthermore, if this issue is not addressed properly, these transitions would affect the final performance of the activity recognition system since they would be misclassified into one of the available classes. Ortiz et al. [14] proposed a transition-aware human activity recognition system for the recognition of six physical activities (walking, walking upstairs, walking downstairs, sitting, standing and lying) and six transitional activities (stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie and lie-to-stand) using the accelerometer embedded in a smartphone. The result showed that the support vector machine could achieve high classification accuracy with a heuristic filtering approach. Noor et al. [15] developed an adaptive segmentation approach for recognizing not only well defined static and dynamic activities, but also transitional activities. In their work, a novel activity transition diagram for activity recognition was developed to validate the activity transition and improve recognition accuracy. Because the samples of posture transitions were, however, much less than basic activities, it would result in a biased classification result. As mentioned previously, the duration of transitions is very short, thus the imbalanced class distribution problem should be taken into consideration during the classification task. Resampling is the most common approach to balance the class distribution in dataset. It includes oversampling the minority class to generate new class samples synthetically, such as SMOTE [16], and under-sampling the majority class by discarding data points. It is noticed that randomly resampling may lead to the possibility of overfitting or useful information loss. Zhang et al. [17] used SMOTE to balance the class distribution in a dataset and set the percentage of data to create for the minority class to 100%. Khoshgoftaar et al. [18] verified that a ratio of 2:1 or even 3:1 in favour of the majority class would result in superior classification performance. In this paper, we considered the transitional activities and the imbalanced class distribution in the dataset, in order to improve the classification accuracy of transitional activities. Our approach is based on an ensemble of heterogeneous classifiers.

Data-driven approaches have been widely used in accelerometer data analysis for activity recognition. These approaches commonly adopted machine learning algorithms and can be categorized into three types: generative approaches, discriminative approaches and heuristic approaches [19]. Generative approaches, such as naïve Bayes [20] and hidden Markov model [21], are flexible, capable of dealing with uncertainty in the data. However, they suffer from the requirement of a large amount of data for training. Discriminative approaches, such as decision tree [22], K-nearest neighbour [23], random forest [24], support vector machine [25] and artificial neural network [26], learn the features mappings to activity labels by creating the decision boundaries in the feature space. For example, decision tree classification models have been successfully adopted for separating static activities from dynamic activities [27]. In terms of dealing with more complex activities, support vector machine and artificial neural network classification methods have been concerned due to their advantages such as robustness in prediction and computational efficiency. However, they face the problem of over-fitting. Heuristic methods use a combination of both, which can generally achieve better performance than any single technique. In addition, recent results in addressing multi-class classification problems have indicated that the adoption of ensembles of classifier models leads to increased classification performance over using only single classifier models [28]. However, much of the previous work on ensembles of classifier models considered only the same type of classifier models [29] [30]. The single classifier has its own strengths as well as weaknesses when dealing with the different classes during the learning process. Moreover, the voting of various algorithms can decrease the bias among the classes occurring in the usage of a single learning algorithm, therefore resulting in a relatively generalized classification [31]. In our experimental study, we focus on a heterogeneous ensemble of classifiers, which consists of classifiers of different types. Our goal is to determine if the heterogeneous ensemble can be used to improve the classification performance by combining various classifiers though the analysis of their classification performance on different activity classes. We have created a dataset with measurements related to different atomic activities (of three types, including transitional activities) and have compared the performance of using isolated classifiers with several topologies of ensemble classifiers in order to validate our approach.

III. Data Acquisition

We summarize in this section the experiment design and data collection process that we have used to obtain the dataset of the three activity types in smart home. Ten healthy adults were instructed to realize twelve activities (see column "Activities" in Table 1) and the data got from an accelerometer placed on their left wrist were collected. These collected data are annotated for the subsequent classification tasks.

Table 1. The taxonomy and description of monitored activities.

Type	Activities	Description
Stationary Activities	Standing	Standing still for 5 minutes
	Sleeping	Sleeping on the sofa for 5 minutes, small movements, such as changing the lying posture, are allowed
	Watching TV	Watching TV while sitting on the sofa in whatever posture the participant feels comfortable for 5 minutes, changing sitting posture is allowed
Dynamic Activities	Walking	Walking on treadmill with a set speed for 5 minutes
	Running	Running on the treadmill for 5 minutes
	Sweeping	Sweeping with the vacuum cleaner in the home area for 5 minutes
Transitional Activities	Stand-to-sit	Standing still for 15s and then sitting on the sofa, repeat for 15 times
	Sit-to-stand	Sitting on the sofa for 10s and then standing up, repeat for 15 times
	Stand-to-walk Walk-to-stand	Performing the "stand-to-walk-to-stand", standing still for 15s then start to walk, keep walking for 15s, then standing still for 15s, repeat for 15 times
	Sit-to-lie	Sitting on the sofa for 15s and then lying down, repeat for 15 times
	Lie-to-sit	Lying on the sofa for 15s and then sitting on the sofa, repeat for 15 times

We aim for the recognition of atomic activities that are necessary to take care of oneself and commonly occur in real daily life. The static activities, such as standing and sleeping, and dynamic activities, such as walking and running, are not limited in duration. Transitional activities, such as stand-to-sit and sit-to-stand, commonly occur within a limited duration, and are characterized by start and end times which usually vary slightly from one person to another [32]. Based on this, static and dynamic activities can be executed continuously, and transitional activities can be executed repeatedly to get separate samples. The twelve activities selected to be recognized are: standing, sleeping, watching TV, walking, running, sweeping, stand-to-sit, sit-to-stand, stand-to-walk, walk-to-stand, lie-to-sit and sit-to-lie. Table

1 shows the classification of these twelve activities into three types and their description. The experiment protocol, data collection and data annotation are described in detail in [7].

IV. METHODS

The acquired acceleration signals from the wrist accelerometer were analysed using MATLAB programs (MathWorks, Natick, MA) The classification task was implemented using Weka [33]. Figure 3 shows the conceptual scheme of the proposed framework to select the most suitable machine learning techniques.

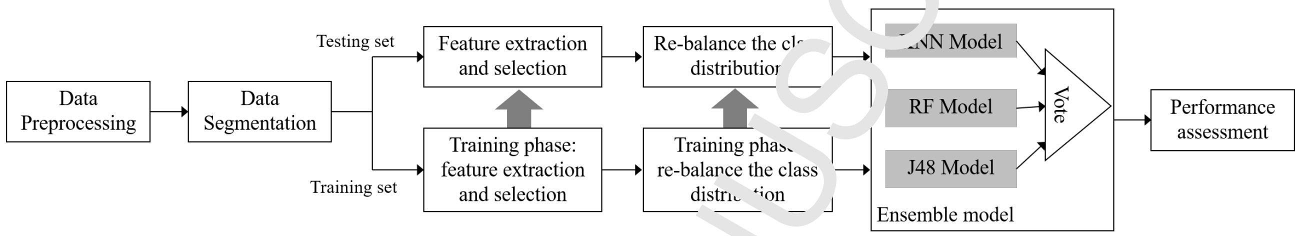


Figure 3. Conceptual scheme of the proposed framework to compare machine learning techniques.

The raw signal data were preprocessed to remove the unwanted noise and separate the body acceleration from gravity acceleration. Then the continuous signals were segmented into windows, each window in the dataset was extracted as an input of 77 features by the feature extraction process. Based on this, the feature selection method was used to select the optimal feature subset by choosing relevant features and removing redundant ones. The resampling technique was applied to rebalance the class distribution in the dataset. After this, we applied a heterogeneous ensemble classification model built by using a majority voting combination function to combine the base classifiers. The computational procedure that we have followed in our experimental setup is illustrated in Figure 4.

4.1. Signal preprocessing

In general, the collected raw data contain signal noise that may be caused by external vibration or loose coupling. Thus we used a third-order low-pass Butterworth filter with cut-off frequency at 20 Hz [34] to eliminate the high frequency noise, and a third order median filter to remove abnormal noise spikes. Moreover, the collected acceleration data were decomposed into the body acceleration (BA) component caused by the body movement, which can be used to distinguish dynamic from static activities, and the gravity acceleration (GA) component caused by the gravity, which can be used to estimate the posture orientation. We used a third-order high-pass Butterworth filter with cut-off frequency at 0.3 Hz to separate the BA and GA components from the filtered signal.

```
Require:
X: Raw acceleration dataset
Y: Raw activity label array
f1(): Low-pass Butterworth filter with 20Hz cut-off frequency
f2(): Median filter
f3(): High-pass Butterworth filter with 0.3Hz cut-off frequency
buffer(): Segmentation function
α(): Feature extraction function
β(): Windows discarding function
μ(): Feature selection function
resampling(): SMOTE with a 1200% boosting oversampling
-----------------------------------------------------------------
procedure SignalPreprocessing(X)
fliteredX = f2(f1(X)); // remove the noise
BX = f3(filteredX); // get the body acceleration
GX = X – BX; //get gravity acceleration
return BX, GX
end procedure
-----------------------------------------------------------------
procedure ObtainFeature(BX, GX, Y)
(SBX, SGX, SY) = buffer(BX, GX, Y); // segment signal into 256 samples
frames with 50% overlapping
A = α(BX,GX); //extract time and frequency domain features
B = mean (Y);
(A', B') = β(A,B); // discard the windows that contain multiple activities
(A_sub, B_sub) = μ(A', B'); // select feature subset
return (A_sub, B_sub)
end procedure
-----------------------------------------------------------------
procedure Classification (A_sub, B_sub)
(x, y) = resampling(A_sub, B_sub); //get balanced features samples
Split (x, y) into training set T_traning and testing set T_testing; //use 10-fold cross
validation
for n = 1 to 10 do
/*Traning*/
  for all (x_i, y_i) ∈ T_traning do
     Build heterogeneous ensemble model C;
  end
/*Testing*/
  for all (x_j, y_j) ∈ T_testing do
     Apply (x_j, y_j) on ensemble model C;
  end
end
return classification accuracy
end procedure
```

Figure 4. Pseudo code of the computational procedure followed in the experimental setup.

4.2. Segmentation and Feature extraction

After the filtering phase, the dataset, which consists of the BA component, the GA component and the activity labels, was divided into windows containing 256 samples each with a 50% overlap between two consecutive windows. Therefore, each classification task made about the activity was performed for the duration of 2.5-second windows. Then the time-domain features and frequency-domain features were obtained by calculating on the values of X-, Y-, and Z-axis of both the BA and the GA components. In addition, the signal magnitude vector (SMV) was used to extract features since it can provide the degree of body movement intensity [35]. The formula is shown in equation (1), where x_t, y_t, z_t refer to the acceleration values of X-, Y-, and Z-axis, respectively, at the sampling time t.

$$\text{SMV} = \sqrt{x_t^2 + y_t^2 + z_t^2} \tag{1}$$

Besides the commonly used time-domain features, such as root mean squared, mean, standard deviation and median absolute deviation [36], the signal magnitude area (SMA), which computed the energy expenditure to distinguish a rest state from a dynamic activity [37], and the tilt angle, which provided information on the participant's orientation, were also included in this work. The frequency-domain features contained spectral entropy, spectral energy, kurtosis, skewness, largest frequency component and signal weighted average. The list of 77 features used in our study as well as their formulations is summarized in Table 2. There are situations where more than one activity label fell into the same segmented window. When dealing with multiple activities within a window, this window was discarded. As a result, the remaining dataset contained 15,178 windows consisting of the extracted features and unique activity labels.

Table 2. Features extracted from time and frequency domain.

Features domain	Features	Formulation	Data values
Time- domain features	Mean	$\frac{1}{n}\sum_{i=1}^{n}$	BA-X, BA-Y, BA-Z, GA-X, GA-Y, GA-Z, BASMV,GASMV
	Root mean squared (RMS)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i}$	
	Standard deviation (STD)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$	
	Median absolute deviation (MAD)	$median(\lvert x_i - median_j(x_j)\rvert)$	
	Range	$\max_i(x)-\min_j(x)$	
	Signal magnitude area (SMA)	$\frac{1}{n}\sum_{j=1}^{m}\sum_{i=1}^{n}\lvert x_i\rvert$	BA, GA, BASMV, GASMV
	Correlation coefficient	$\mathrm{cov}(x,y)/\sigma_x\sigma_y$	BA-XY, BA-YZ, BA-XZ, GA-XY, GA-YZ, GA-XZ,
	Tilt angle (TA)	$\tan^{-1}(x/\sqrt{y^2+z^2})*180/\pi$	BA-X, BA-Y, BA-Z
Frequency-domain features	Spectral energy	$\frac{1}{n}\sum_{i=1}^{n}x_i^2$	BA-X, BA-Y, BA-Z, BASMV
	Spectral entropy	$\sum_{i=1}^{n}(c_i\log(c_i)), c_i = s_i/\sum_{j=1}^{n}s_j$	
	Skewness	$E\left[(\frac{x-\bar{x}}{\sigma})^3\right]$	
	Kurtosis	$E[(x-\bar{x})^4]/E[(x-\bar{x})^2]^2$	
	Largest frequency component	$\arg\max_i(x_i)$	
	Frequency signal weighted average	$\sum_{i=1}^{n}(ix_i)/\sum_{j=1}^{n}x_j$	

4.3. Feature selection

Feature selection aims to identify a subset of the most discriminative features that can increase the classification performance, as well as remove the redundant features that contribute no additional information to the classifier. Previous research mainly used two main categories of feature selection methods: filter methods and wrapper methods [38]. In this experimental study, three filter feature selection methods were considered: Information Gain (InfoGain), correlation-based feature selection (CFS) with Genetic search (GS) algorithm and ReliffF. These three filter methods were selected since the filter methods for feature selection depend on general data characteristics rather than predetermined classifiers [39]. Since the ensemble classifier approach that we aimed to assess is based on heterogeneous classifiers, wrapper methods were not suitable in our case.

4.4. Class distribution rebalance

Having an imbalanced number of activity instances may lead to inconsistencies in the activity recognition model. Previous researches have demonstrated that randomly over-sampling techniques would lead to over-fitting problems and under-sampling techniques would result in the loss of useful information [40]. In this study, we utilized SMOTE boosting method to improve the representation of classes with less number of instances, which generates synthetic instances along the segments adding any of the k minority class to their nearest neighbours. With an oversampling rate of 1200% boosting to increase the samples of the minority classes, 15,517 new samples were synthesized to balance minority and majority classes.

4.5. Classification methods

For our experimental work, we considered and assessed several classification techniques in two steps. The first step was to generate a diverse and accurate set of base classifiers by comparing the performance of various single classifiers. In this study, we considered six commonly used classifiers, namely: random forest (RF), k-nearest neighbours (KNN), decision tree (J48), artificial neural network (MLP), Naïve Bayes (NB) and support vector machine (SVM) [41]. These six classifiers belong to different classifier types, which have different internal representations and may be biased in different ways. The different outputs of the base classifiers represented the extent to which they disagree about the probability distribution for the test data. This diversity would lead to the disagreement with each other over the data instances covering a range of the feature space.

The second step was to combine the models into a heterogeneous set of base classifiers using combination functions. The combination functions intend to make the best use of the information obtained from the base classifiers for the purpose of making class labels predictions as accurately as possible. There are two general categories of combination functions: fusion and selection. Since the output of the selection function typically depends on the characteristics of the instance whose class is being predicted, it is not suitable in this study. Thus we chose the majority voting function [42], which is a fusion function, to combine the outputs of the base classifiers. In majority voting, the outputs from base classifiers are used as votes on the predicted class, and the prediction voted the most is outputted as the final predicted class of the ensemble classifier. The majority voting function is defined as shown in equation (2).

$$\mathrm{C(x)} = \mathrm{argmax} \sum_{C_{i(x)=y}} 1 \quad (2)$$

As a result, 4 base classifiers (2 RF, 1 KNN and 1 J48) were selected to generate the ensemble. The predictions of these four classifiers, which were the results of 3 algorithms applied on 12 balanced class distribution dataset, were finally voted to get the final classification result.

V. RESULTS AND DISCUSSION

In this section, we present the results we have obtained in our experimental setup by applying different feature selection methods, a resampling method and various learning algorithms on the dataset. The main performance indicators, such as accuracy, sensitivity and specificity, are based on the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) figures obtained by executing the corresponding algorithms in Weka. This section shows with a specific example how the followed methodology is valid to select a set of suitable machine learning technologies and detect a set of atomic activities, based on the measurements gathered by the smart object.

5.1. Performance assessment of feature selection methods

In order to determine the optimal feature subset which contains the most relevant features, we compared the effectiveness of the 10-fold cross validation performance of three feature selection methods (Table 3). The performance obtained when using all features was used as baseline for comparison. The decision tree (J48) was used here as learning algorithm for classification. We used three measures to validate the performance of the three feature selection methods: accuracy (defined as (TP+TN)/(TP+TN+FP+FN)), sensitivity (defined as TP/(TP+FN)), and specificity (defined as TN/(TN+FP)).

Table 3. Comparison of the performance on five feature subsets by three feature selection methods.

Feature subset	No. of features	Accuracy	Sensitivity	Specificity
All features	77	88.98%	0.890	0.990
FS1: InfoGain	75	89.21%	0.892	0.990
FS2: ReliefF	40	89.32%	0.893	0.990
FS3: CFS+GS	36	89.34%	0.893	0.990

In FS1 and FS2, the features were ranked according to InfoGain and ReliefF, respectively. The top 75 ranked features by InfoGain yielded an accuracy of 89.21% and a sensitivity of 0.892. The top 40 ranked features by ReliefF produced an accuracy of 89.32% and a sensitivity of 0.893. The CFS and GS algorithm (FS3) directly selected a feature subset containing 36 features and provided the highest accuracy of 89.34% and the highest sensitivity of 0.893 among the three feature subsets. Compared to the performance of using all features, the classifiers with less features, in contrast, produced higher accuracy and sensitivity. This is probably due to the fact that some features are irrelevant and redundant, and their high correlation with other features would affect the classification performance. As a result, the CFS and GS algorithm was chosen as the feature selection method to reduce the feature vector dimensions as well as to improve the classifier performance. The 36 features picked by the CFS and GS algorithm are given in Table 4. These selected features were used as input to various learning algorithms in further data analysis phases.

Table 4. The features selected by CFS and GS algorithm.

Feature domain	Selected features
Time-domain	Mean-BA-X, STD-BA-X, MAD-BA-X, MAD-BA-X, SMA-BA, TA-BA-X, TA-BA-Y, TA-BA-Y, Mean-GA-X, Mean-GA-Y, Mean-GA-Z, RMS-GA-X, RMS-GA-Y, RMS-GA-Z, MAD-GA-Z, Range-GA-Y, Range-GA-Z, SMA-GA, Mean-BASMV, RMS- BASMV, SMA-BASMV, RMS-GASMV, STD-GASMV, MAD-GASMV, Range-GASMV, SMA-GASMV
Frequency-domain	Spectral Entropy-BA-Z, Skewness-BA-X, Skewness-BA-Y, Kurtosis-BA-Y, Largest frequency component-BA-X, Frequency signal weighted average-BA-Y, Frequency signal weighted average-BA-Y, Spectral Entropy-BASMV, Skewness-BASMV, Kurtosis-BASMV

5.2. Validation of base classifiers performance

In order to choose the most efficient base classifiers (that will be also used for the construction of an ensemble of classifiers later), we investigated and compared various classification algorithms, namely: RF, KNN, J48, MLP, NB and SVM. The preliminary activity classification performance of six different algorithms on the balanced dataset with selected features is summarized in Table 5. The data were processed using average 10-fold cross validation during the training and testing phases.

For the performance validation methods, besides the accuracy, sensitivity and specificity, we also introduced the following measures: mean absolute error (MAE), relative absolute error (RAE), F-measure and receiver operating characteristic (ROC) area. The mean absolute error (MAE) is a quantity used to measure how close the predictions are to the eventual outcomes. The definition of MAE is given by equation (3).

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1,n}|y - y'| \qquad (3)$$

The relative absolute error (RAE) represents the error as a percentage of the true value and measures how far the predictions are from the eventual outcomes. The definition of RAE is given by equation (4).

$$\mathrm{RAE} = \sum_{i=1,n}|y - y'| / \sum_{i=1,n}|\overline{y'} - y'| \qquad (4)$$

Where y is the actual value, y' is the predicted value and n is the number of samples. The ideal case corresponds to both MAE and RAE being 0. F-measure is defined as 2*Sensitivity*(1-Specificity) / (1-Specificity+Sensitivity). The receiver operating characteristic (ROC) curve is a graphic plot of the true positive rate against false positive rate at various threshold settings. Therefore the area under the ROC curve can be used to illustrate the performance of a classifier.

Table 5. Comparison of the performance of different learning algorithms.

Algorithms	Accuracy	MAE	RAE	Sensitivity	Specificity	F-measure	ROC Area
Random forest (RF)	96.09%	0.025	16.01%	0.961	0.997	0.961	0.999
K-nearest neighbours (KNN)	95.64%	0.007	4.80%	0.956	0.996	0.956	0.976
Decision tree (J48)	89.34%	0.019	12.41%	0.893	0.990	0.893	0.953
Artificial neural network (MLP)	85.97%	0.028	18.09%	0.860	0.987	0.859	0.976
Naïve Bayes (NB)	63.21%	0.063	41.27%	0.632	0.967	0.616	0.933
Support vector machine (SVM)	31.86%	0.114	74.33%	0.681	0.938	0.319	0.628

Among the six classifiers, RF, KNN and J48 showed their superior performance in evaluation metrics. Specifically, RF contributed the highest accuracy of 96.09%. It also achieved the best performance in terms of sensitivity, specificity, F-measure and ROC area measures of 0.961, 0.997, 0.961 and 0.999, respectively. This is because RF combines multiple decision trees with various rules, which can better handle the multi-class data characteristics. Thus, we used RF as the baseline classifier here. Compared to RF, the KNN yielded lower MAE and RAE measurements of 0.007 and 4.80%, respectively. This means that KNN produced closer predictions to the eventual outcomes. The J48 classifier also contributed lower MAE (0.019) and RAE (12.41%) values than RF, whereas KNN and J48 performed better in the classification of specific activity classes. Compared to the performance of RF, the classification results from MLP, NB and SVM were not very good. The accuracy obtained by MLP was 85.97% and SVM only contributed 31.86% accuracy. Therefore, we considered RF, KNN and J48 as the possible ensemble members.

To determine the best classification performance for each activity, we analysed the confusion matrices of the twelve activities produced by RF, KNN and J48 (see Table 6). We found that these three classifiers made quite different misclassifications. The matrices revealed that the RF model often misclassified walking class (A4) to sweeping (79 instances), stand-to-walk (35 instances), walk-to-stand (24 instances) and sit-to-lie (22 instances). Standing class (A1) was frequently confused with walk-to-stand (50 instances). Sweeping class (A6) was sometimes misclassified to walk (25 instances). Likely these activities shared similar movements. J48, for its part, contributed less misclassified instances of standing (A1), watching TV (A3) and walking (A4) activities compared to RF. KNN also contributed less

Table 6. The confusion matrices of the twelve activities obtained by the three selected learning algorithms.

Random Forest (RF)

	ST	SL	WT	WA	RU	SW	ST-SI	SI-ST	ST-WA	WA-ST	LI-SI	SI-LI
A1	2290	0	3	0	0	16	0	6	9	50	2	3
A2	5	2348	0	0	0	6	7	2	1	1	4	5
A3	0	0	2302	2	0	3	8	3	4	10	11	28
A4	0	1	0	2071	1	79	9	17	95	24	8	22
A5	0	0	0	4	2270	7	0	0	2	0	0	0
A6	2	0	0	25	2	2233	11	17	17	18	10	6
A7	0	0	0	1	0	1	2279	31	4	9	2	0
A8	0	0	1	0	0	1	18	2350	1	6	0	2
A9	1	0	0	24	0	14	7	2	2331	27	8	4
A10	4	0	2	7	0	6	3	17	46	2289	3	2
A11	0	3	0	0	0	0	5	3	0	2	2338	15
A12	0	1	0	0	0	0	1	2	0	0	26	2375

k-Nearest Neighbours (KNN)

	ST	SL	WT	WA	RU	SW	ST-SI	SI-ST	ST-WA	WA-ST	LI-SI	SI-LI
A1	2264	2	6	9	0	10	0	6	23	56	3	0
A2	1	2353	10	1	0	2	5	0	0	0	1	6
A3	1	3	2330	6	0	0	6	3	1	2	8	11
A4	2	0	3	2151	0	39	9	11	71	25	6	10
A5	0	0	0	4	2276	3	0	0	0	0	0	0
A6	3	0	0	60	1	2179	15	17	40	15	7	4
A7	0	0	0	3	0	8	2240	61	3	7	2	3
A8	1	1	0	0	0	2	46	2316	1	8	4	0
A9	11	0	1	40	1	19	7	10	2225	99	5	0
A10	100	4	3	13	0	4	4	13	115	2120	1	2
A11	0	6	0	0	0	0	8	1	1	6	2310	34
A12	0	3	5	0	0	0	3	3	0	0	37	2354

Decision Tree (J48)

	ST	SL	WT	WA	RU	SW	ST-SI	SI-ST	ST-WA	WA-ST	LI-SI	SI-LI
A1	2281	0	2	4	0	20	4	9	8	47	1	3
A2	1	2332	6	0	0	5	6	1	2	10	4	12
A3	1	6	2279	6	0	8	13	7	1	20	13	17
A4	5	2	6	1962	5	75	8	20	172	47	15	10
A5	0	0	0	4	2266	8	1	0	4	0	0	0
A6	16	4	4	79	6	2056	28	37	55	35	11	10
A7	1	11	12	11	0	20	2022	161	18	26	21	24
A8	4	3	5	21	0	22	148	2061	35	42	19	19
A9	7	1	3	144	2	63	23	30	1995	129	9	12
A10	40	17	20	51	1	33	35	51	145	1959	19	8
A11	1	23	1	16	0	10	27	15	10	15	2034	214
A12	5	13	17	16	2	7	26	22	5	8	199	2085

A1 = ST = Standing, A2 = SL = Sleeping, A3 = WT = Watching TV, A4 = WA = Walking, A5 = RU= Running, A6 = SW= Sweeping, A7 = ST-SI= Stand-to-sit, A8 = SI-ST= Sit-to-stand, A9 = ST-WA= Stand-to-walk, A10 = WA-ST= Walk-to-stand, A11 = LI-SI=Lie-to-sit, A12 = SI-LI= Sit-to-lie.

misclassified instances than RF of several activities. For example, RF misclassified 79 instances of walking to sweeping, whereas KNN only misclassified 39 instances of walking to sweeping. The results illustrated that no single classifier performs best in all evaluation measures (even if they may be the best in a subset of the evaluation measures).

Thus combing these three single classifiers may contribute to a better classification performance in our experiment. Based on this result, we selected RF, KNN and J48 as ensemble members due to their excellent performance in evaluation measures. These three classifiers had their own strengths on classifying different activities, they also provided different disagreements with each other, and this disagreement was cleared by majority voting to obtain a final prediction. As a result, we defined an ensemble model with 4 heterogeneous classifiers (2 RF, 1KNN and 1J48).

5.3. Comparison of performance of our ensemble with imbalanced and all features datasets

We evaluated the classification performance of the ensemble constructed with 2 RF, 1 KNN and 1 J48 classifiers by applying it on the balanced feature subset. Table 7 presents the results, which showed high sensitivity and specificity (sensitivity ranged from 0.893 to 0.995, specificity from 0.993 to 1.000) on each activity class, which indicates that the ensemble model performed good both in positive and negative identifications. The ensemble model performed extremely well in identifying running with the highest F-measure of 0.997. Activities such as sleeping, watching TV, lie-to-sit and sit-to-lie were also identified very well. Compared to the single classifiers, the ensemble model provided 97.00% accuracy, higher than the 96.09% accuracy obtained by RF alone, which was the highest accuracy among the various single classifiers. These results indicate that the ensemble model is a high-performance method to achieve relatively higher accuracy than the single classifiers.

Table 7. The performance of heterogeneous ensemble model with selected features.

	ST	SL	WT	WA	RU	SW	ST-SI	SI-ST	ST-WA	WA-ST	LI-SI	SI-LI
Sensitivity	0.965	0.987	0.972	0.893	0.995	0.956	0.979	0.988	0.964	0.962	0.989	0.989
Specificity	1.000	1.000	1.000	0.998	1.000	0.995	0.998	0.996	0.993	0.995	0.997	0.997
F-measure	0.980	0.992	0.984	[illegible]	0.997	0.950	0.976	0.975	0.946	0.952	0.981	0.977

ST = Standing, SL = Sleeping, WT = Watching TV, WA= Walking, RU= Running, SW= Sweeping, ST-SI= Stand-to-sit, SI-ST= Sit-to-stand, ST-WA= Stand-to-walk, WA-ST= Walk-to-stand, LI-SI= Lie-to-sit, SI-LI= Sit-to-lie.

To evaluate the performance of the constructed ensemble classifier, we derived and compared three different ensemble models: EM1, EM2 and EM3. EM1 was derived on the raw data without applying the resampling technique, all features extracted from the raw data were used as the input. EM2 used all features trained on the balanced data after performing the resampling techniques. Thus the difference between EM1 and EM2 was the class distribution in dataset. EM3 was derived on the balanced data using the 36 selected features obtained by CFS and GS algorithm. The difference between EM2 and EM3 was the number of used features for building the ensemble models.

Table 8 presents the performance of EM1, EM2 and EM3 in terms of accuracy, sensitivity, specificity, F-measure and ROC area. Comparing the performance of EM1 and EM2, the accuracy of EM2 was increased from 95.14% to 96.01%. The sensitivity of EM2 was increased to 0.960, which was higher than that of EM1 (0.951). Even though the accuracy of EM1 was still relatively high, many instances of transitional activities were misclassified, this can be shown by the sensitivity of EM1 for each activity (see Figure 5). This is perhaps due to the learning algorithms assuming that positive and negative classes are balanced in datasets, thus they usually maximize the overall accuracy by a bias towards the majority class. To address the poor performance on minority classes, we applied the SMOTE with an oversampling rate of 1200% boosting to balance the dataset.

Table 8. The performance of different ensemble models generated by 10-fold cross validation.

Dataset	Accuracy	Sensitivity	Specificity	F-measure	ROC Area
EM1: Imbalanced data with all features	95.14%	0.951	0.993	0.946	0.972
EM2: Balanced data with all features	96.01%	0.960	0.996	0.960	0.978
EM3:Balanced data with 36 selected features	97.00%	0.970	0.997	0.970	0.984

Figure 5 depicts the sensitivity of EM1, EM2 and EM3 on each class. It shows that after rebalancing the dataset, the sensitivity of EM2 and EM3 on transitional activity classes was significantly increased and balanced with the static and dynamic activity classes. By applying the feature selection method to obtain a feature subset that contains only discriminative and relevant features, compared to EM2, the performance of EM3 was improved from 96.01% to 97.00% in terms of accuracy, from 0.960 to 0.970 in terms of sensitivity, from 0.996 to 0.997 in terms of specificity, from 0.960 to 0.970 in terms of F-measure and from 0.978 to 0.984 in terms of ROC area.

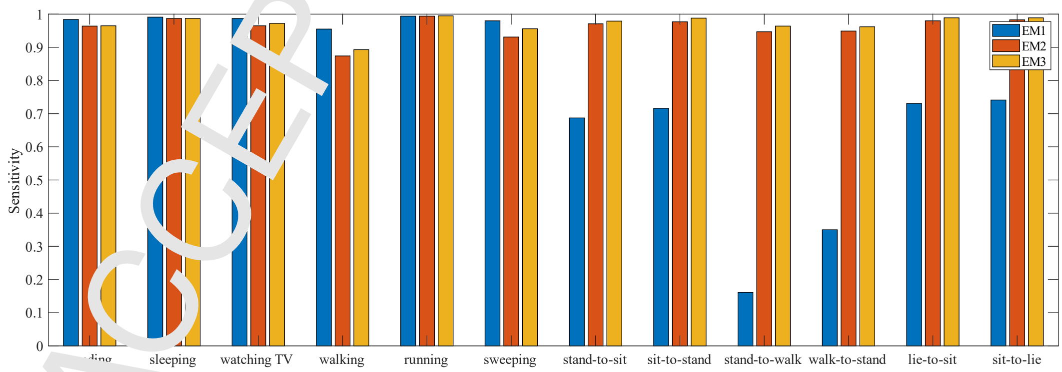


Figure 5. Comparison of sensitivity of the twelve activities between EM1, EM2 and EM3.

The results showed that not only combining the heterogeneous base classifiers by majority voting, but also rebalancing class distribution by resampling technique as well as selecting informative and relevant features by feature selection method, have the ability to improve the performance of activity recognition.

5.4. Performance of our ensemble applied to publicly available datasets

Previous results relate to our 10-hour dataset obtained as described in section I[illegible] above. In this section, we show the results of our ensemble of heterogeneous classifiers applied to three public datasets, namely: PAMAP2 [43], SBHAR [44] and UoA [14]. Whereas SBHAR and UoA contain samples of transitional activities, PAMAP2 does not. Table 9 shows the accuracy obtained per activity by the ensemble classifier applied to each of these three datasets.

Table 9. Comparison of accuracy of activity recognition of our proposal when applied to different published datasets.

Dataset	ST	SL	WT	WA	RU	SW	ST-SI	SI-ST	ST-WA	WA-ST	LI-SI	SI-LI
PAMAP2 [47]	0.989	0.961	NA	0.940	0.987	NA	NA	NA	NA	NA	NA	NA
SBHAR [48]	0.976	0.951	NA	0.924	NA	NA	[illegible]	0.949	NA	NA	0.928	0.926
UoA [14]	0.938	NA	NA	0.940	NA	NA	0.876	0.852	NA	NA	0.920	0.912
Collected dataset	0.965	0.987	0.972	0.893	0.995	[illegible]	0.979	0.988	0.964	0.962	0.989	0.989

ST = Standing, SL = Sleeping, WT = Watching TV, WA= Walking, RU= Running, SW= Sweeping, ST-SI= Stand-to-sit, SI-ST= Sit-to-stand, ST-WA= Stand-to-walk, WA-ST= Walk-to-stand, LI-SI=Lie-to-sit, SI-LI= Sit-to-lie

The classification performance of static and dynamic activities on the three public datasets all contributed higher than 90% in terms of recognition accuracy. On some occasions, the accuracy is even higher for the other three datasets than for the one collected by us. Moreover, the recognition accuracy of transitional activities of our ensemble was kept similarly high when applied to other datasets different and independent from ours (slightly lower, but always higher than 85% and most of the times above 90%). With these results we show that the high performance of our selected ensemble of heterogeneous classifiers is not completely dependent on the specific dataset that contains the activities' samples. This, together with the facts that our participants were not instructed on how to perform the activities and that the public datasets referenced in this section have been provided by three different research teams who are totally independent from ours, makes us confident that our proposed method to select the classifiers, filters and oversampling is useful to accurately detect the considered three types of activities with a reasonable independence on how different people may perform them.

VI. Discussion and Conclusion

In this paper we have described the steps followed for the design of the decision-making subsystem associated to a simple and non-intrusive device to measure and classify atomic activities at home. We consider an architecture in which the detection of these atomic activities are provided by the IoT infrastructure present in home. From these atomic activities, Ambient Intelligence (AmI) techniques can infer, taking into account the contextual information also gathered by sensors, more complex (or composite) activities that can be related to the inhabitants' behaviour.

The set of possible atomic activities is not fully defined and can evolve with the technology and habits of the population. In addition, the technology to detect the activities can be very heterogeneous and it is even possible to have several IoT technologies obtaining data to detect the same set of activities. We have proposed the use of a smart object consisting on a 3D accelerometer embedded in a wrist-worn device. We have provided with the specification of a complete procedure for the gathering, mining and cleaning of the raw data, rebalancing of the different classes' number of samples and comparison of diverse classification procedures, to end with the selection and assessment of an ensemble of heterogeneous classifiers.

Our experiments showed good classification accuracy among twelve atomic activities that cannot be overlapped in time because of their mutually exclusive nature. These results were attained by the oversampling of minority classes together with the selected ensemble of three different classifiers whose results are combined using a majority voting function. One important contribution is related to the demonstration of the complete procedure that allows to compare different possible machine learning methods to classify the measurements coming from a simple and easily accepted wrist-worn IoT smart object. Depending on the population sector and the specificities of the activities that are most important to be detected, this same procedure is applicable to select the most suitable classification technique.

On the other hand, by reviewing the existing literature we have observed that most of the proposed activity detection solutions are not well adjusted for the different types of ambulatory activities. In fact, the so-called transitional activities present a lower detection rate than the rest, due to characteristics such as their short duration and difficulty to be predicted. This lower detection rate causes a higher imprecision when it comes to infer a person's behaviour, which in turn may lead to wrong decisions being made afterwards.

In this work we have proved the feasibility of detecting a set of twelve activities belonging to three types, including short-timed transitional activities, by using a single non-intrusive data acquisition device together with a data processing mechanism consisting on cleaning and balancing the sample set and using ensemble supervised-learning classifiers. The empirical results show how, even though the detection precision we have obtained is moderately lower for some static and dynamic activities, it is higher for transitional activities up to values comparable to the other types of activities. We have got similarly good results with our selected data processing and classification method applied to three publicly available data sets different from ours. Thus, we conclude that our proposed solution is able to attain high success rates and, equally importantly, shows a balanced performance throughout the three activity types investigated.

With respect to the specific device type used to measure the data, i.e. the 3D wrist-worn accelerometer, we consider it as an object that is easily accepted among the majority of the population. In fact, many people already use watch-like devices and bracelets on a regular basis, and they have proven very usable. Besides, the fact that it is personally assigned to each person makes it easier to identify each individual and thus to support multi-user data. It is worthy to note that the wrist-worn sensor has limitations to monitor some lower limb-related activities, such as cycling, thus the collaboration with sensors placed on other parts of the body, such as chest and ankle, is planned to be included in the future work. On the downside of this election, we are aware that, since all our experiments have included only one sensor modality on the non-intrusive device, any additional contextual information (such as heart rate and location) has been left out. The addition of context information to the output of the ensemble classifiers is something worth investigating to assess the possible benefits on both the detection precision and the tolerance to classification errors, and gives also interesting additional input to the AmI subsystem to detect complex activities and behaviours. Another important issue to be tackled is the assessment of the precision that we would obtain under actual real life conditions. In fact, all presented experiments contained activities that were adequately timed so that the classifiers did not face windows containing mixed activities samples. Our hypothesis on this is that it is unlikely that the number of non-homogeneous windows (i.e. windows containing samples from different activities) in a real-life deployment would be significant with respect to the number of homogeneous windows. This would facilitate to filter out non-homogeneous windows and thus minimize their distortional effects.

REFERENCES

1. Lloret, J.; Canovas, A.; Sendra, S.; Parra, L. A smart communication architecture for ambient assisted living. IEEE Communication Magzine. 2015(53): 26–33.

2. Roy, P.C.; Abidi, S.R.; Abidi, S.S.R. Possibilistic activity recognition with uncertain observations to support medication adherence in an assisted ambient living setting. Knowledge Based System. 2017(133):156–173.
3. Singh, A.; Bianchi-Berthouze, N.; Williams, A.C.C. Supporting everyday function in chronic pain using wearable technology. In Proceedings of the Conference on Human Factors in Computing Systems, Denver, USA, 6–11 May 2017: 3903–3915.
4. Ghayvat, H.; Mukhopadhyay, S.; Shenjie, B.; Chouhan, A.; Chen, W. Smart home based ambient assisted living: Recognition of anomaly in the activity of daily living for an elderly living alone. In Proceedings of the IEEE International Instrumentation and Measurement Technology Conference, Houston, TX, USA, May 2018: 14–17.
5. Meditskos, G.; Plans, P.M.; Stavropoulos, T.G.; Multi-modal activity recognition from egocentric vision, semantic enrichment and lifelogging applications for the care of dementia. Journal of Visual Communication and Image Representation. 2018(51): 169-190.
6. Ni, Q.; García Hernando, A.B.; de la Cruz, I.P. The elderly's independent living in smart homes: A characterization of activities and sensing infrastructure survey to facilitate services development. Sensors, 2015(15): 11312–11362.
7. Ni, Q.; Cleland I.; Nugent C.D.; García Hernando, A.B.; de la Cruz, I.P. Dataset obtained for detecting static, dynamic and transitional activities of daily life by using a single wrist-worn accelerometer. Data in Brief. Submitted.
8. Chernbumroong, S.; Cang, S.; Atkins, A.; Yu, H. Elderly activities recognition and classification for applications in assisted living. Expert Systems with Applications, 2013(40): 1662-1674.
9. Wang, Z.; Wu, D.; Chen, J.; Ghoneim, A.; Hossain, M.A. A triaxial accelerometer-based human activity recognition via eemd-based features and game-theory-based feature selection. IEEE Sensors Journal, 2016(16): 3198-3207.
10. Gao, L.; Bourke, A.K.; Nelson, J. Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems. Medical engineering & physics, 2014(36): 779-785.
11. Cleland, I.; Kikhia, B.; Nugent, C.; Boytsov, A.; Hallberg, J.; Synnes, K.; McClean, S.; Finlay, D. Optimal placement of accelerometers for the detection of everyday activities. Sensors, 2013(13): 9183-9200.

12. Mehrang S, Pietilä J, Korhonen I. An Activity Recognition Framework Deploying the Random Forest Classifier and A Single Optical Heart Rate Monitoring and Triaxial Accelerometer Wrist-Band. Sensors, 2018, 18(2): 613.
13. Qi J, Yang P, Hanneghan M, et al. A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors. IEEE Internet of Things Journal, 2018. DOI 10.1109/JIOT.2018.2846359.
14. Reyes-Ortiz, J.-L.; Oneto, L.; Samà, A.; Parra, X.; Anguita, D. Transition-aware human activity recognition using smartphones. Neurocomputing 2016(171): 754-767.
15. Noor, M.H.M.; Salcic, Z.; Kevin, I.; Wang, K. Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer. Pervasive and Mobile Computing, 2017(38): 41-59.
16. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. Journal of artificial intelligence research, 2002(16): 321-357.
17. Zhang, S.; McClean, S.I.; Nugent, C.D.; Donnelly, M.P.; Galway, L.; Scotney, B.W.; Cleland, I. A predictive model for assistive technology adoption for people with dementia. IEEE Journal of Biomedical and Health Informatics, 2014(18): 375-383.
18. Khoshgoftaar, T.M.; Seiffert, C.; Hulse, J.V.; Napolitano, A.; Folleco, A. Learning with limited minority class data. Sixth International Conference on Machine Learning and Applications, Cincinnati, USA, 13-15 Dec, 2007: 348-353.
19. Chen, L.; Hoey, J.; Nugent, C.D.; Cook, D.J.; Yu, Z. Sensor-based activity recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews. 2012(42): 790-808.
20. Liu L, Wang S, Su G, et al. Towards complex activity recognition using a Bayesian network-based probabilistic generative framework. Pattern Recognition. 2017(68): 295-309.
21. Safi, K.; Mohammed, S.; Attal, F.; Khalil, M.; Amirat, Y. Recognition of different daily living activities using hidden markov model regression. Third Middle East Conference on Biomedical Engineering, Beirut, Lebanon, 6-7 Oct, 2016: 16-19.
22. Hemminki, S.; Nurmi, P.; Tarkoma, S. In Accelerometer-based transportation mode detection on smartphones, Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, ACM, 2013: 13.
23. Ren, X.; Ding, W.; Crouter, S.E.; Mu, Y.; Xie, R. Activity recognition and intensity estimation in youth from accelerometer data aided by machine learning. Applied Intelligence, 2016(45): 512-529.

24. Thomaz E, Essa I, Abowd G D. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2015:1029-1040.
25. Mazumder, O.; Kundu, A.S.; Lenka, P.K.; Bhaumik, S. Ambulatory activity classification with dendogram-based support vector machine: Application in lower-limb active exoskeleton. Gait & Posture, 2016(50): 53-59.
26. Ronao, C.A.; Cho, S.-B. Human activity recognition with smartphone sensors using deep learning neural networks. Expert Systems with Applications, 2016(59): 235-244.
27. Chernbumroong, S.; Atkins, A.S.; Yu, H. In Activity classification using a single wrist-worn accelerometer, Software, Knowledge Information, Industrial Management and Applications (SKIMA), 2011 5th International Conference on, 2011; IEEE: 1-6.
28. Daghistani, T.; Alshammari, R. Improving accelerometer-based activity recognition by using ensemble of classifiers. International Jouornal of Advanced Computer Science and Applications. 2016(7): 128-133.
29. Dietterich, T.G. Ensemble methods in machine learning. International workshop on multiple classifier systems, Springer, Berlin, Heidelberg, 2000: 1-15.
30. Jurek, A.; Nugent, C.; Bi, Y.; Wu, S. Clustering-based ensemble learning for activity recognition in smart homes. Sensors. 2014(14): 12285-12304.
31. Krawczyk B. Active and adaptive ensemble learning for online activity recognition from data streams. Knowledge-Based Systems. 2017(138): 69-78.
32. Reyes-Ortiz, J.-L.; Oneto, L.; Ghio, A.; Samá, A.; Anguita, D.; Parra, X. Human activity recognition on smartphones with awareness of basic activities and postural transitions. International Conference on Artificial Neural Networks. Springer, Cham, 2014: 177-184.
33. Garner, S.R. In Weka: The waikato environment for knowledge analysis, Proceedings of the New Zealand computer science research students conference, 1995: 57-64.
34. Karantonis, D.M.; Narayanan, M.R.; Mathie, M.; Lovell, N.H.; Celler, B.G. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Transactions on Information Technology in Biomedicine. 2006(10): 156-167.
35. Cheng, W.-C.; Jhan, D.-M. Triaxial accelerometer-based fall detection method using a self-constructing cascade-adaboost-svm classifier. IEEE Journal of Biomedical and Health Informatics. 2013(17): 411-419.

36. Figo, D.; Diniz, P.C.; Ferreira, D.R.; Cardoso, J.M. Preprocessing techniques for context recognition from accelerometer data. Personal and Ubiquitous Computing. 2010(14): 645-662.
37. Mathie, M.; Celler, B.G.; Lovell, N.H.; Coster, A. Classification of basic daily movements using a triaxial accelerometer. Medical and Biological Engineering and Computing. 2004(42): 679-687.
38. Zhang, M.; Sawchuk, A.A. In A feature selection-based framework for human activity recognition using wearable multimodal sensors, Proceedings of the 6th International Conference on Body Area Networks. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011: 92-98.
39. Liu, H.; Yu, L. Toward integrating feature selection algorithms for classification and clustering. IEEE Transactions on Knowledge and Data Engineering. 2005(17): 491-502.
40. Chawla, N.V. Data mining for imbalanced datasets: An overview. In Data mining and knowledge discovery handbook, Springer, 2005: 853-867.
41. Wu, X.; Kumar, V.; Quinlan, J.R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Philip, S.Y. Top 10 algorithms in data mining. Knowledge and Information Systems. 2008(14): 1-37.
42. Bagui, S.C. Combining pattern classifiers: Methods and algorithms. Technometrics. 2005(47): 517-518.
43. O. Banos, M. Damas, H. Pomares, I. Rojas, M.A. Tóth, O. Amft, A benchmark dataset to evaluate sensor displacement in activity recognition. ACM Conference on Ubiquitous Computing, 2012: 1026-1035.
44. D. Anguita, A. Ghio, L. Oneto, X. Parra, J.-L. Reyes-Ortiz, A public domain dataset for human activity recognition using smartphones, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2013.