A preliminary study for automatic activity labelling on an elder people ADL dataset

Enrique de la Cal[1], Mirko Fáñez[2], Alvaro DaSilva[2], Jose Ramón Villar[1], Javier Sedano[2], and Victor Suárez[3]

[1] University of Oviedo, Computer Science Department,
Oviedo, Spain
{delacal,villarjose}@uniovi.es
[2] Instituto Tecnológico de Castilla y León, Pol. Ind. Villalonquejar,
09001, Burgos, Spain
mirko@mirkoo.es, javier.sedano@itcl.es, ada@ubu.es
[3] University of Oviedo, Control and Automatica Department,
EPI, Gijón, Spain
vmsuarez@uniovi.es

Abstract. One consequence of the aging population is an increase in life expectancy implying greater healthcare needs as well as a serious healthy aging program. So healthy aging is one of the main challenges in the first world nowadays, and as much as possible devices, software and technological solutions applied to measure and improve the quality of life of the elder people are necessary.
Recently, we presented a first prototype of an activity monitoring kit, and this study includes the analysis of the dataset gathered after six months of use. Since the wearable devices employed in this monitoring kit have not the automatic activity recognition service available, current work proposes several techniques to label automatically the Time Series (TS) obtained in the experiment. Thus, a new device with the same sensors as the old one plus the automatic activity recognition service available will be used to obtain a new labelled dataset, that will be used to learn a new model using semi-supervised learning to tag the not-labelled dataset.

Keywords: ADL automatic identification · Falls in Elderly · Wearable sensors · Fall Detection · Human Activity Recognition

1 Introduction and motivation

The old-age dependency ratio (people aged 65 and above relative to those aged 15 to 64) in the EU is projected to increase by 21.6 percentage points, from 29.6% in 2016 to 51.2% in 2070. This implies that the EU would go from having 3.3 working-age people for every person aged over 65 years to only two working-age people [2]. One consequence of the ageing population is an increase in life expectancy implying greater healthcare needs [3]. Thus, non-invasive tools to monitor and analyse Elderly health and activity are required. The most common non-invasive and easy-to-use tools to measure the activity in elderly are the wearable devices. Static daily living activities like standing, sitting and lying are simple to detect, whilst dynamic ones such as walking, running, jumping, are more difficult to recognise. Concerning this, two classification techniques have been regularly used: the supervised and unsupervised activity classification approaches. Traditionally the supervised activity classification approaches have not been considered in real free-living environments where external factors can affect negatively their performances. Moreover, the collection of sufficient amounts of labelled data for a representative set of free-living activities may be sometimes difficult to achieve and computationally expensive. On the other hand, unsupervised machine learning represents the second approach used in Human Activity Recognition (HAR). In this case, labelled data is not required which can overcome the aforementioned limitations of the supervised techniques [5]. There are two categories of unsupervised recognition models: Static [4] and temporal classification approaches [6].

In previous work a first prototype of an autonomous, low-cost and easy-to-use elderly activity monitoring kit was presented [1]. This kit includes a set of 12 smart-bands (2 TICWATCH E2 and 10 SAMSUNG Gear Fit2) with 3DACC, GYROSCOPE and HR Sensors, as well as other components (miniPC and 4G router) to store and access the data remotely. This prototype was deployed in a nursery house in June 2019, and it has been gathering data for 6 months.

While the prototype presented [1] had the main goal of gathering data to obtain a dataset with real falls, current work will be focused on the analysis of the levels of daily life activity excluding Falls. It is worth to state that the obtained data has been split in two datasets: i) one labelled dataset obtained from two participants wearing TICWATCH smartwatch, and ii) other no-labelled dataset gathered from the 10 participants wearing a SAMSUNG smart-band. Thus, a simple automatic labelling technique based on semi-supervised learning is proposed to label the SAMSUNG dataset using a classification model obtained from the TICWATCH dataset. Furthermore, in order to contrast the labelling results, a statistical study involving several movement features like AOM, ACC and SMA was included.

This work is structured as follows: next section includes the design issues of the semi-supervised proposal presented here while the experimentation and the discussion of the results are coped in section 3. Finally, conclusions and future work is included in section 4.

2 The proposal

The main goal of our proposal is to analyse and characterize the daily levels of activity of the unlabelled data collected for 6 months from a group of participants using the activity monitoring kit presented in [1].

As, this first prototype of monitoring kit used a model of smart-band (OLD-DEVICES) with the automatic activity identification service not available, two units of these smart-bands were replaced by other two new models of smartwatch (NEWDEVICES) with this capability activated. The OLDDEVICES use the following sensors: a 3D Accelerometer, a gyroscope and a heart rate sensor, whilst the NEWDEVICES have the same sensors plus the automatic activity identification service activated. The NEWDEVICES have been collecting data for the last 2 months of the experiment and they will replace all the OLDDEVICES in the next release of the monitoring kit.

Hence, the idea is to learn a model of activity level labelling using the NEWDEVICES and use semi-supervised learning to apply these models to the OLDDEVICES dataset in order to label the activity of the participants.

Consequently, it's proposed a method based on the following steps: i) OLDDEVICES and NEWDEVICES datasets clean and pre-processing, and ii) Design and perform an automatic segmentation algorithm taking as input the NEWDEVICES dataset, and deploy the models on the OLDDEVICES dataset.

2.1 OLDDEVICES dataset clean and pre-processing

The big volume of data obtained for 6 months needs to be pre-processed and cleaned since some days either the participants did not wear the monitoring device or several OLDDEVICES ran out of battery quickly because of an operating system failure. Thus, several statistics have been considered to remove the waste data:

- Mean (MEAN_p) and Standard Deviation (STDN_p) of the number of hours recorded by day for participant p.
- Hours per Day Threshold (HDT_p): Minimum number of recorded hours to consider a day valid for the participant p. HDT_p=0.7*MEAN_p
- Percentage of Recorded Days (PoRD_p): The percentage of recorded days out of the 6 months, for the participant p.
- Valid Percentage of Recorded Days (VPoRD_p): The percentage of recorded days out of the 6 months, for the participant p with a number of hours overpassing HDT_p.

Therefore, all the days with a number of recorded hours under HDT_p, as well as all the data of those participants with a VPoRD_p under 30%, will be removed. This subsection will be covered later on the Numerical Results section (see 3.2).

2.2 Automatic segmentation of Activities of the OLDDEVICES dataset

This study proposes to label the activity level of the OLDDEVICES dataset segmenting the TSs in high and low activity periods.

We have decided to define an algorithm based on the HR sensor to segment the TSs in high and low activity. Therefore, a simple algorithm based on thresholds is defined:

1. Select the TS windows on the NEWDEVICES dataset that has been automatically labelled by the Android Activity Recognition API (using sliding

windows of 10 seconds) as ON_FOOT series (walking or running and labelled as HIGH) and STILL series (no activity or low activity and labelled as LOW).
2. Calculate the mean HR on both types of TSs, grouped by participant (ONFOOT_HR$_p$ and STILL_HR$_p$).
3. Calculate the mean HR on both types of TSs (not by participant, ONFOOT_MEAN and STILL_MEAN).

Accordingly, the activity level threshold is calculated as:

$$ONFOOT\_TH = ONFOOT\_MEAN - 0.7 * (ONFOOT\_MEAN - STILL\_MEAN) \quad (1)$$

So the TS windows with HR lower than ONFOOT_TH is considered Low-Activity; High-Activity otherwise.

Finally, the ONFOOT_TH will be deployed on the OLDDEVICES dataset to segment the TS windows, correlating the type of activity with different well-known features on HAR related with the intensity of movement like Simple Moving Averages (SMA), Amount Of Movement (AOM) and Magnitude of Acceleration (MAG).

3 Numerical results

3.1 Materials and methods

The devices: concerning the specific brand and model of the OLDDEVICES and NEWDEVICES referred above, we can say that for the experiments included in this section we have considered the smart-band SAMSUNG Gear Fit 2 as the OLDDEVICES model, and the smartwatch TICWATCH E2 as the NEWDEVICES model (see Fig. 1).

The participants When the first prototype of monitoring kit was presented, it was defined a very strict protocol of participant inclusion and exclusion supervised by an expert gerontologist [1]. As a product of this protocol a group of 10 people with ages between 76 and 98 was recruited.

The timeline of the experiment: the experiment lasted six months between June 2019 and November 2019, collecting data from the 10 participants using the SAMSUNG devices, and from the 10th October to 22th November two SAMSUNG devices were replaced by other two TICWATCH ones for participants #1 and #2.

Fig. 1: Monitoring kit release 0.0 with SAMSUNG devices #1 and #2 replaced by TICWATCH devices.

The methods: The first stage is the clean and pre-processing of the OLDDEVICES and NEWDEVICES datasets. After this stage, the OLDDEVICES dataset will be analysed by performing a semi-supervised learning using the HR threshold-based (HRT) models learned with the NEWDEVICES dataset to label the OLDDEVICES dataset. Hence, this section comprise the following steps: i) both datasets will be cleaned and pre-processed, ii) the HRT will be estimated on the NEWDEVICES dataset in order to segment the datasets in High-Activity and Low-Activity TSs and iii) finally the OLDDEVICES dataset will be characterized analyzing the segmentation based on the HRT.

Facilities and running time: The experiments were carried out on a 2.4GHz Intel Core i9 with 32GB of RAM MACOSX Laptop. With this configuration, the most time-consuming R script was the OLDDEVICES TSs Characterization based on the HRT values learned with the NEWDEVICES dataset (the last step), which took 6 hours to complete (not using R parallel execution).

3.2 Dataset clean and pre-processing

Due to erratic battery behaviour for the 6 months that data was being recorded with the OLDDEVICES, the dataset was not very homogeneous among all participants. This leads to a data consolidation process. Table 1 shows data statistics previous to the consolidation, which will be used to perform this pre-processing stage. In the light of these results, we must establish as a valid day threshold the mean recorded hours per day (HDT, calculated as mentioned in section 2.1), and use it to calculate the percentage of valid recorded days per participant

(VPoRD). There are 2 participants (#2 and #9) that has low VPoRD, which will not be considered in this study, keeping the rest of them (although not all have the same number of valid recorded days, there is a consistent minimum).

Table 1: Results for both the un-consolidated features (Registered hours per day, HDT, PoRD) and the consolidated feature VPoRD (after applying the HDT), for the OLDDEVICES dataset.

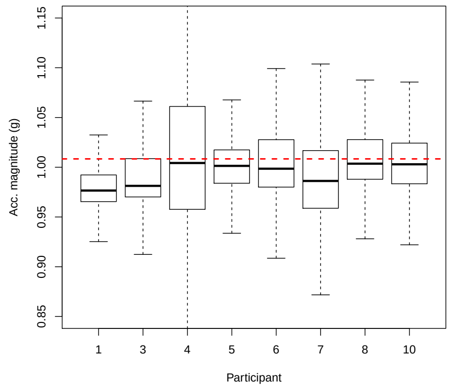
PartID	**Regist. hours/day**		**HDT (hours)**	**PoRD (%)**	**VPoRD (%)**
	Mean	**Std**			
01	7.1173	1.6168	4.9821	38.70	38.70
02	6.9631	0.3994	4.8742	3.22	**3.22**
03	7.4811	1.1911	5.2368	40.32	37.09
04	7.0347	2.2072	4.9242	33.87	32.25
05	6.9846	0.0712	4.8892	32.25	32.25
06	6.6304	0.7328	4.6413	90.32	87.09
07	8.4405	2.2326	5.9084	37.09	35.48
08	7.0163	0.9142	4.9114	90.32	87.09
09	6.4032	1.6624	4.4822	17.74	**16.12**
10	6.8757	0.9580	4.8130	87.09	80.64
mean	**7.0947**	**1.1986**	**4.9663**	**47.0920**	**44.9930**
std	**0.5520**	**0.7259**	**0.3864**	**31.1535**	**29.6479**

3.3 Automatic segmentation of activities.

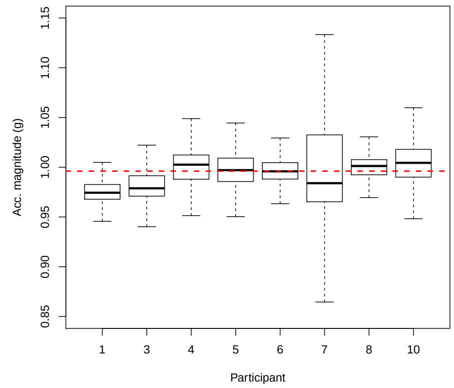
Table 2 shows the ONFOOT_TH and STILL_TH statistics computed using the data of the two participants belonging to the NEWDEVICES dataset. The ONFOOT_TH HR threshold is calculated based on the equation 1:

$$ONFOOT\_TH = 69.7261 \tag{2}$$

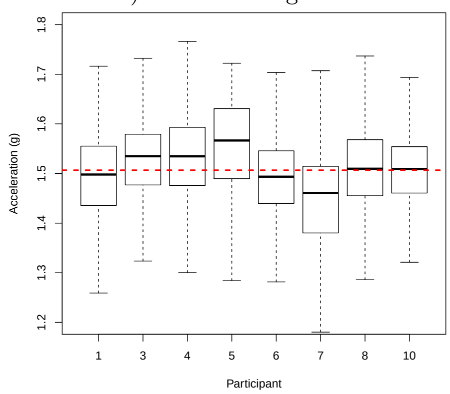
Once the ONFOOT_TH is calculated and deployed on the OLDDEVICES dataset, the TSs of this dataset are classified as High-Activity and Low-Activity TSs. In order to characterize these two activity levels, the well-known HAR features: Acceleration Magnitude, SMA and AOM were calculated. Figure 2 includes the boxplots of these features arranged in two columns: first column includes the High activity boxplots while second column includes the Low activity ones. The results show that AOM and Acceleration Magnitude standard deviation are decisive in this characterization, since its values are totally correlated with the activity level. The mean (red dashed line) as well as the deviation of the AOM High Activity TS (c-i) is clearly higher than the corresponding for Low Activity TSs(c-ii). Regarding the other features, just the deviation of the ACCMAG is clearly correlated with the level of activity: the deviation of the Low Activity TSs (a-ii) is significantly lower than the High Activity TSs (a-i). Table 3 shows the numerical values corresponding to the boxplots included in Figure 2.



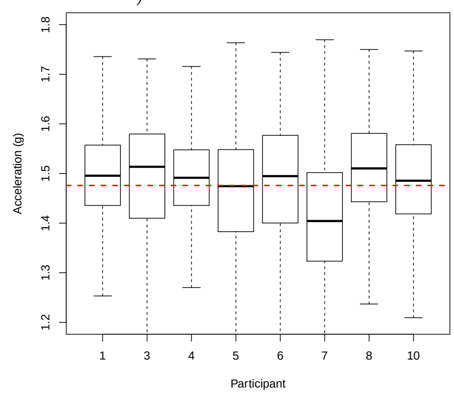
a-i)ACCMAG High-act



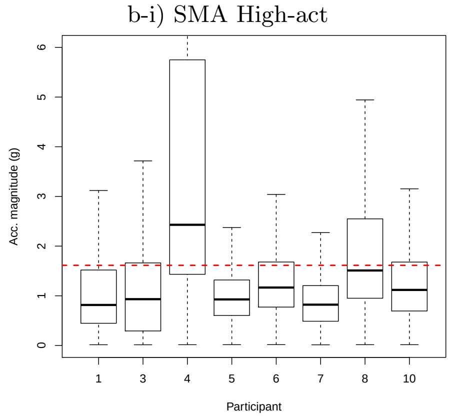
a-ii)ACCMAG Low-act



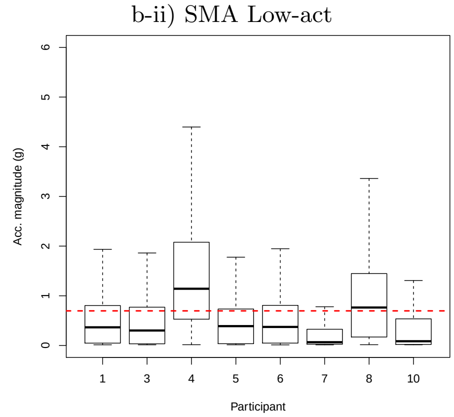
b-i) SMA High-act



b-ii) SMA Low-act



c-i) AOM High-act



c-ii) AOM Low-act

Fig. 2: Boxplots for the 8 participants considered in this study, segmented by activity level: High-Act and Low-Act. The red dashed line is the mean value for all the participants.

Table 2: HR segmentation results for both participants in the NEWDEVICES dataset.

PartID	**ONFOOT**		**STILL**	
	mean	**std**	**mean**	**std**
1	88.5301	22.77587	56.74898	16.42893
2	89.5713	24.79433	66.13932	11.46430
mean	**89.0507**	**23.7851**	**61.4442**	**13.9466**
std	**0.7362**	**1.4273**	**6.6400**	**3.5105**

Table 3: Results for OLDDEVICES dataset, after segmentation of data in High-Activity and Low-Activity, using ONFOOT_TH HR threshold learned with the NEWDEVICES dataset.

pID	**High-Activity TS**						**Low-activity TS**					
	Acc. magn.		**SMA**		**AOM**		**Acc. magn.**		**SMA**		**AOM**	
	Mean	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**
01	0.9839	0.1187	1.4919	0.0928	1.2328	1.3764	0.9774	0.0564	1.4938	0.0961	0.6203	0.8378
03	0.9932	0.1177	1.5112	0.1212	1.1845	1.2105	0.9839	0.0540	1.4798	0.1515	0.5096	0.6556
04	1.0352	0.3105	1.5337	0.0877	3.5741	2.6082	1.0092	0.1444	1.4885	0.0966	1.6632	1.7859
05	1.0044	0.0816	1.5475	0.1088	1.0244	0.7686	0.9984	0.0484	1.4551	0.1330	0.4875	0.5956
06	1.0092	0.1235	1.4929	0.0836	1.3898	1.0554	0.9977	0.0513	1.4786	0.1303	0.5607	0.7205
07	0.9911	0.0930	1.4446	0.1071	0.9736	0.8440	0.9935	0.0395	1.4190	0.1604	0.2255	0.3427
08	1.0386	0.2831	1.5287	0.1551	2.2347	2.1323	1.0046	0.0970	1.5075	0.1073	1.1643	1.5425
10	1.0115	0.1159	1.5041	0.0841	1.2896	0.9036	1.0042	0.0363	1.4846	0.1109	0.3411	0.5377
mean	**1.0084**	**0.1555**	**1.5068**	**0.1050**	**1.6129**	**1.3624**	**0.9961**	**0.0659**	**1.4759**	**0.1233**	**0.6965**	**0.8773**
std	**0.0200**	**0.0887**	**0.0320**	**0.0243**	**0.8840**	**0.6648**	**0.0108**	**0.0367**	**0.0274**	**0.0244**	**0.4787**	**0.5104**

4 Conclusion and future work

This study presents a method to characterize the activity levels of a real not-labelled ADL dataset based on a semi-supervised automatic labelling technique using the HR sensor. The automatic segmentation model has been learned taking as input the automatically labelled dataset gathered from the TICWATCH E2 smartwatches. This model has been deployed on a not-labelled long-term dataset collected using SAMSUNG Gear Fit 2 smart-bands. The results obtained state that the automatic segmentation model based on HR classify quite coherently the TS of the SAMSUNG dataset in High and Low levels of activity. In addition, the AOM feature and the standard deviation of the Acceleration magnitude show a high correlation with the level of activity, verifying that the classification is quite good.

Considering that the baseline of this study was a Faller Monitoring kit [1] that has been collecting data for 6 months without a valid fall, we think that the experiment has been successful. So, next release of the monitoring kit will comprise TICWATCH smartwatches instead of SAMSUNG Gear Fit 2 smart-bands, since the first ones are more robust and stable.

Acknowledgement

This research has been funded partially by Spanish Ministry of Economy, Industry and Competitiveness (MINECO) under grant TIN2017-84804-R and by Foundation for the Promotion of Applied Scientific Research and Technology in Asturias, under grant FC-GRUPIN-IDI2018000226.

References

1. de la Cal, E., DaSilva, A., Fez, M., Villar, J., Sedano, J., Surez, V.: An autonomous fallers monitoring kit: release 0.0. In: Proceedings of the 19th International Conference on Intelligent Systems Design and Applications (2019)
2. Comission, E.: 2018 Ageing Report: Policy challenges for ageing societies (2020 (accessed February 12, 2020)), `https://ec.europa.eu/info/news/economy-finance/policy-implications-ageing-examined-new-report-2018-may-25_en`
3. King, R.C., Villeneuve, E., White, R.J., Sherratt, R.S., Holderbaum, W., Harwin, W.S.: Application of data fusion techniques and technologies for wearable health monitoring. Medical Engineering and Physics **42**, 1 – 12 (2017)
4. Quante, M., Kaplan, E.R., Rueschman, M., Cailler, M., Buxton, O.M., Redline, S.: Practical considerations in using accelerometers to assess physical activity, sedentary behavior, and sleep. Sleep health **1**(4), 275–284 (2015)
5. Trabelsi, D., Mohammed, S., Amirat, Y., Oukhellou, L.: Activity recognition using body mounted sensors: An unsupervised learning based approach. In: The 2012 International Joint Conference on Neural Networks (IJCNN). pp. 1–7. IEEE (2012)
6. Trabelsi, D., Mohammed, S., Chamroukhi, F., Oukhellou, L., Amirat, Y.: An unsupervised approach for automatic activity recognition based on hidden markov model regression. IEEE Transactions on automation science and engineering **10**(3), 829–835 (2013)