Using Ensembles for Improving Fall Detection

1st José R. Villar
Computer Science Department
University of Oviedo
Oviedo, Spain
villarjose@uniovi.es

2nd Enrique de la Cal
Computer Science Department
University of Oviedo
Oviedo, Spain
delacal@uniovi.es

3rd Víctor M. González
Electrical Engineering Department
University of Oviedo
Gijón, Spain
vmsuarez@uniovi.es

4thJavier Sedano
Applied Electronics and Artificial Intelligence Group
Instituto Tecnológico de Castilla y León
Burgos, Spain
javier.sedano@itcl.es

***Abstract*—Technology can highly enhance the quality of life of elderly people when living autonomously. Actually, fall detection is one of the possible improvements: having an oracle to request help in case of falling may succed in provining confidence to adults who have already fallen, for instance. In a recent study, a wrist worn wearable solution has been proposed based on threshold for peak detection and several low computational models for classifying the peak events as related with falls. It was found that ensembles could lead to a better performance of the system. In this research, a preliminary study in ensembling the classifiers is proposed, analysing the outcome in each case for each of the different scenarios initially proposed. Although a very simple ensemble solution is used, results suggest this may be the solution when a robust fall detection system is needed.**

***Index Terms*—Fall Detection, Neural Networks, Support Vector Machines, Ensemble of Classifiers**

I. Introduction

Fall Detection (FD) is a very active research area, with many applications in health care, work safety, etc. [1]. Even though there are plenty of commercial products, the best rated products only reach 80% of success [2]. There are basically two types of FD systems: context-aware systems and wearable devices [3], [4]. FD has been widely studied using context-aware systems, i.e. video systems [5]; nevertheless, the use of wearable devices is crucial because of the high percentage of elderly people and their desire to live autonomously at their own house [6].

By far, tri-axial accelerometry (3DACC) is the most used option within the literature of FD [7]–[11]. Different solutions have been proposed to perform the FD, for instance, a feature extraction stage and SVM have been applied directly in [7], [9], using some transformations and thresholds with very simple rules for classifying an event as a fall [10]–[12]. A comparison of classifiers has been presented in [8].

As stated in [13], the most common location of the 3DACC sensor is on the waist. This location has been proven valid for FD when the subject suffers from a impairment disease; however, it might not be the best option when focusing on autonomous population. It is suggested that, for this sort of people, a wrist-based solution would suit better. Furthermore, in [13], an improvement of a FD method has been proposed using a 3DACC wearable device placed on a wrist. Although good results were given, the solution needs further improvements to reduce the false alarms or the undetected events. One of the possible solutions is to use an ensemble of classifiers using those models that performed the best. In this study, a very simple combination of these methods is proposed, analysing the outcome of the ensemble of the methods.

Next section deals with the description of the FD system and the details of the model ensemble. Section III deals with the datasets and the evaluation of the hypothesis. Section IV shows and discusses the obtained results. Finally, conclusions are drawn.

This research has been funded by the Spanish Ministry of Science and Innovation, under projects MINECO-TIN2014-56967-R and MINECO-TIN2017-84804-R.

II. Fall Detection with Wrist-worn Wearables

The method was originally proposed in [14], being extended in [13]. Basically, it is a threshold-based peak detection stage followed by feature extraction and classification stages. Several different models were proposed. In this study we suggest ensembling them to improve the FD.

In the next subsection, the peak detection and feature extraction stages are detailed, while in subsection II-B the different models and the ensemble of the outcome are described.

A. Peak detection and feature extraction

As proposed in [13], [14], a very simple finite state machine is used to detect the falls; this state machine is shown in Figure 1. The data gathered from a 3DACC located on the wrist is processed using a sliding window. A peak detection is performed, and if a peak is found, the data within the sliding window is analysed to extract several features which are ultimately classified as FALL or NOT_FALL. The FD block is performed with a classifier. In [15] it was claimed that the lower the computational cost of the classifier the better as it must be run in the wearable device.

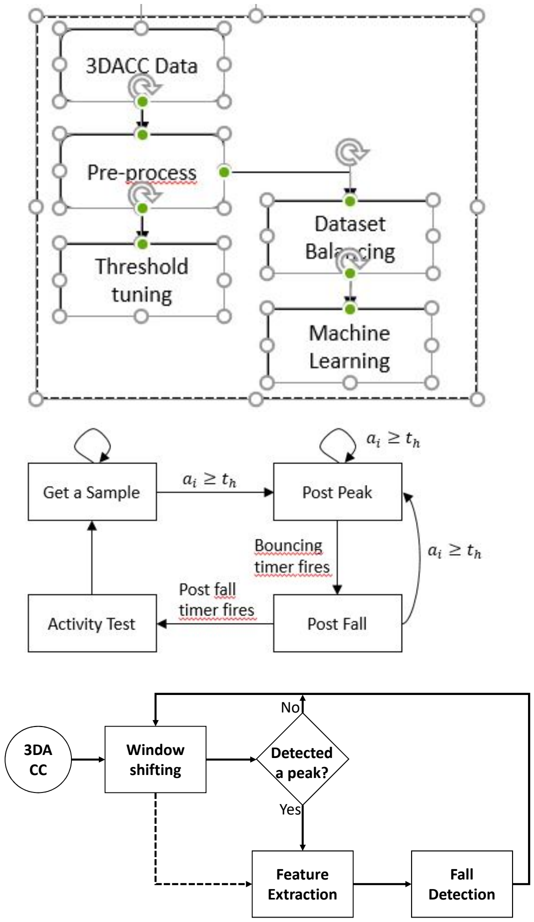

Fig. 1: Block diagram of the solution. Upper part: both the threshold and the classifier are trained using a dataset. Center part: the finite state machine, the bouncing and post fall timers are set to 1000 and 1500 ms, respectively. Bottom part: the features are computed and the extracted sample is classified whenever a peak is detected.

The feature extraction is executed whenever a peak is detected and follows the dynamics within a fall -refer to Fig. 2-. Let us assume that gravity be $g = 9.8m/s$. Given the current timestamp t, we find a peak at **peak time** $pt = t - 2500ms$ (point 1) if at time pt the magnitude of the acceleration a_t -see Equation 1- is higher than $th_1 = 3 \times g$ and there is no other peak in the period $(t - 2500ms, t]$ (no other a value higher than th_1). If this condition holds, then it is stated that a peak occurred at pt.

$$a_t = \sqrt{a_{tx}^2 + a_{ty}^2 + a_{tz}^2} \tag{1}$$

The **impact end** (ie) (point 2) denotes the end of the fall event; it is the last time for which the a value is higher than $th_2 = 1.5 \times g$. Finally, the **impact start** (is) (point 3) denotes the starting time of the fall event, computed as the time of the first sequence of an $a <= th_3$ ($th_3 = 0.8 \times g$) followed by a value of $a >= th_2$. The impact start must belong to the interval $[ie - 1200\ ms, pt]$. If no impact end is found, then it is fixed to $pt + 1000$ ms. If no impact start is found, it is fixed to pt. From now on and without loosing generalization, as long as we we know the sampling frequency, we can refer to timestamp or to positions within a sliding window that includes the samples in $[is, ie]$.

When using subindex i we refer to the sample position within the sliding window, and when using subindex t we refer to a timestamp; however, they are interchangeable. When a peak is detected the feature extraction is performed, computing for this peak time several parameters and features.

With these three times -is, pt, ie- calculated, the following transformations should be computed:

- Average Absolute Acceleration Magnitude Variation, $AAMV = \sum_{t=is}^{ie} \frac{|a_{t+1} - a_t|}{N}$, with N the number of samples in the interval.
- Impact Duration Index, $IDI = ie - is$.
- Maximum Peak Index, $MPI = max_{t \in [is, ie]}(a_t)$.
- Minimum Valley Index, $MVI = min_{t \in [is-500, ie]}(a_t)$.
- Peak Duration Index, $PDI = pe - ps$, with ps the peak start defined as the time of the last magnitude sample below $th_{PDI} = 1.8 \times g$ occurred before pt, and pe, the peak end defined as the time of the first magnitude sample below $th_{PDI} = 1.8 \times g$ occurred after pt.
- Activity Ratio Index, ARI, calculated as the ratio between the number of samples not in $[th_{ARIlow} 0.85 \times g, th_{ARIhigh} = 1.3 \times g]$ and the total number of samples in the 700 ms interval centered in $(is + ie)/2$.
- Free Fall Index, FFI, the average magnitude in the interval $[t_{FFI}, pt]$. The value of t_{FFI} is the time between the first acceleration magnitude below $th_{FFI} = 0.8 \times g$ occurring up to 200 ms before pt; if not found, it is set to $pt - 200\ ms$.
- Step Count Index, SCI, measured as the number of peaks in the interval $[pt - 2200, pt]$.

As stated in the Introduction section of this study, several solutions for FD are based on threshold to detect peaks plus an extra processing [10], [16]–[18]. The solution proposed in [14] is not different. Furthermore, several thresholds are used in that study not only to detect a peak but also to compute the extracted features. All of them have been fixed by analysing the gathered data, establishing some typical values for the features for the class FALL. In otder to obtain a better threshold, [13] proposed the use of computational intelligence and optimization, suggesting that Genetic Algorithms and Simulated Annealing would find most suitable values for the threshold.

B. Ensemble of models

According to the block diagram, each sample of these eight features is classified as a fall event or not using the predefined model. Therefore, this model has to be trained; this topic is covered in the next subsection.

As explained in [13], several different low computational models were proposed to classify each peak as Fall or Not_Fall; the constraint of reduced computational capacity is a current limit in nowadays wearable devices. Therefore,

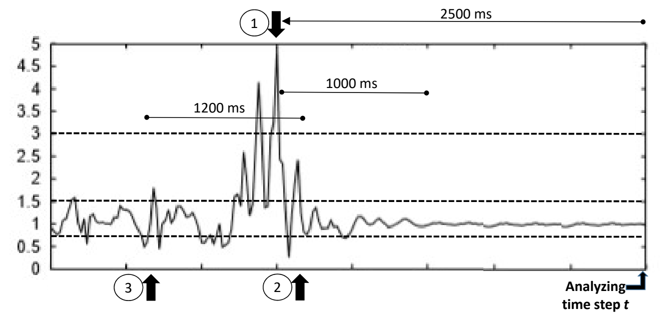

Fig. 2: Graph elaborated from [14], showing the evolution of the magnitude of the acceleration in multiples of g. Analysing the signal at time stamp ***t***, the peak condition described in the text must be found in order to detect a fall. The x-axis represents the time; each mark corresponds to 500 ms.

Decision Trees (DT) suing C5.0, Rule Based Systems (RBS) also using C5.0, and Support Vector Machines (SVM) were added to the original option of feed-forward Neural Networks (NN).

Provided there exists a collection of Time Series (TS) with data gathered from real falls or from Activities of Daily Living (ADL), a training phase can be proposed to train the FD model. Let us consider a dataset containing $\{TS_i^L\}$, with $i = 1 \cdots N$, where n is the number of TS samples and L is the assigned label; that is, a sample of this dataset is a TS_i^L with the data gathered from a participant using a 3DACC on the chosen location, i.e., on a wrist. Let us assume we know a priori whether this TS_i^L includes or not the signal gathered when a fall occurred; therefore, each TS is labelled as $L = FALL$ or $L = NOT\_FALL$.

Now, let us evaluate the peak detection and the feature extraction blocks for each TS. Whenever a TS_i^L has no peak, the TS_i^L is discarded. When a peak is detected for TS_i^L, then the eight features are computed, and label L can be assigned to this new sample. Therefore, a new dataset is created with M eight features labelled samples, with $M \leq N$. This dataset was used in [14] to train the feed-forward NN.

Nevertheless, it has been found that this solution i) might generate more than a sample for a single TS_i^L -which is not a problem-, and ii) certainly will generate a very biased dataset, with the majority of the samples belonging to the class FALL. From their study [14], it can be easily drawn that the main reason of a 100% of detection is this biased dataset.

Consequently, in this research we propose to include a dataset balancing stage using SMOTE [19], so at least a $40/60$ ratio is obtained for the minority class.

Therefore, in this study we discard the method of DT because it was the one that was found with higher number of undetected events and higher number of false alarms. The remaining three models (RBS, SVM and NN) will be used as independent classifiers; the ensemble is computed as the weigthed sum of the outcomes of the three models, using the same weight for each one. As mentioned before, this is a preliminary study and just a simple solution is fetched in order to evaluate whether the outcome of the system can be enhanced or not.

III. Material and Methods

A. Public Datasets

A common way of studying FD is by developing a dataset of simulated falls plus extra sessions of different ADL. All these TS are labelled and become the test set for the corresponding study. The vast majority of these datasets were gathered with the sensor attached to the main body -either on the chest, waist, lumbar area, or thigh-.

In this research, three publicly available 3DACC datasets are used:

- UMA Fall dataset [20] includes data gathered from 3DACC sensors placed on different parts of the body - ankle, waist, wrist and, head- while performing simulated falls; this is the type of data needed in this research because the main hypothesis of this study is to perform FD with a sensor worn on a wrist. Furthermore, there is no pattern in the number of repetitions of each activity or fall simulation. Some participants did not simulate any fall, some performed 6 or 9, and one participant simulated 60 falls.
- UNIOVI dataset [21] includes ADL and simulated epileptic seizures [22]. This one has been considered because it includes a high movement activity -the simulated partial tonic-clonic seizures- followed by a relatively calm period plus some other ADL, all of them measured using 3DACC placed on the dominant wrist.
- DaLiaC dataset [23] includes several sensors, one on the wrist and one on the waist among others. Up to 19 young healthy participants and up to 13 different ADL are considered, from sitting to cycling.

B. The Experimentation Scheme

As proposed in [13], a participant based cross validation (cv) scheme is performed, including training, testing and validation. Once a participant is chosen for validation -or train and test-, all the TS gathered for that participant are included in the validation -or train and test- dataset. The 15% of the participants from the UMA Fall and the UNIOVI datasets have been chosen for validation, the remaining participants are assigned to the training and testing dataset.

The general process is depicted in Fig. 3. The training and testing dataset was used in [13] for tuning the threshold to perform the peak detection; three thresholds were considered: 2.5 as the minimum value of a peak for the 3DACC magnitude, 3.0 given in [14] and 3.09590 obtained from the optimization stage.

The peak detection algorithm is run on the TS belonging to the train and test dataset; for each peak detected, the features are extracted and labelled as FALL or NOT_FALL. This procedure produces a feature extraction dataset; this dataset is used to train the models using either 5x2 cv or 10-fold cv. In this research, 5x2 cv has been used. Beacuse this dataset might be highly imbalanced, SMOTE is applied to obtain a more

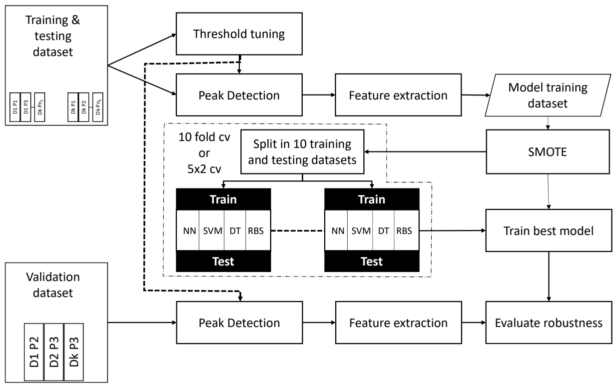

Fig. 3: The Machine Learning process within the cross validation scheme. The training and testing dataset is used for i) threshold optimization, and ii) peak detection and feature extraction. The labelled dataset is then used for the machine learning process to find the best modelling option. The best option is then evaluated with the validation dataset once processed so the real performance of the system can be obtained.

suitable dataset to use in the learning process; the balanced dataset must have a minority class percentage within 40%-60%. Afterwards, the models are learnt using grid search for the best parameter subset.

Finally, the validation dataset is considered. It goes through the peak detection block -using the optimized threshold- and, whenever a peak is found, the feature extraction stage is executed. Finally, the eight features are classified using the best model found in the previous stage. A TS from the validation dataset will be classified as FALL if a peak is detected and the subsequent classifier outputs the FALL label; otherwise, the TS will be assigned the label NO_FALL.

In this study, the same participant based cv and the same models used in [13] are proposed. Only the validation is performed differently. This is because the hypothesis is that using ensembles of the models obtained in the train and test stage can improve the overall performance of the FD system. Therefore, only the validation stage is performed, comparing the results of each individual model with the very simple ensemble scheme detailed before.

IV. Experimentation's Results

A. Obtained Results

Table I and Table II show the results obtained in the validation; the former includes the confusion matrices, while the latter shows the standard statistical measurements. Although the number of TS belonging to the Fall class are relatively too small, it seems that the ensemble obtained performance is nearly as good as the best model in terms of the number of detected events, while at the same time reduces the number of false alarms. Still, some undetected fall events can be observed, though.

TABLE I: Confusion matrices for the three analysed thresholds and for each model type: feed-forward NN, rule base systems learned with C5.0 (RBS) and, Support Vector Machines (SVM). ENS stands for ensembling the outcome of the models.

Threshold 2.5					
	Reference			**Reference**	
NN	**Fall**	**Not Fall**	**RBS**	**Fall**	**Not Fall**
Fall	10	47	**Fall**	10	42
Not Fall	2	250	**Not Fall**	2	245
SVM	**Fall**	**Not Fall**	**ENS**	**Fall**	**Not Fall**
Fall	8	18	**Fall**	10	37
Not Fall	4	279	**Not Fall**	2	260
Threshold 3.0					
	Reference			**Reference**	
NN	**Fall**	**Not Fall**	**RBS**	**Fall**	**Not Fall**
Fall	12	52	**Fall**	11	29
Not Fall	0	245	**Not Fall**	1	268
	Reference			**Reference**	
SVM	**Fall**	**Not Fall**	**ENS**	**Fall**	**Not Fall**
Fall	10	12	**Fall**	11	29
Not Fall	2	285	**Not Fall**	1	268
Threshold 3.09290					
	Reference			**Reference**	
NN	**Fall**	**Not Fall**	**RBS**	**Fall**	**Not Fall**
Fall	12	59	**Fall**	12	35
Not Fall	0	238	**Not Fall**	0	262
	Reference			**Reference**	
SVM	**Fall**	**Not Fall**	**ENS**	**Fall**	**Not Fall**
Fall	10	13	**Fall**	12	31
Not Fall	2	284	**Not Fall**	0	266

B. Discussion

Clearly, the number of TS in the validation dataset were not enough to gather a proper conclusion. It is simply too small to evaluate the general performance. Therefore, the first issue to solve is the number of TS. In a recent study [24], up to twelve publicly available datasets related with FD and ADL were compared. However, these datasets show as a very sparse effort, with different ADLs, goals, population, etc. besides, the UNIOVI dataset [21] also includes ADLs. Moreover, each of published datasets includes its own set of 3DACC sensors, placed on different body locations. Some of them include only data from the waist, others include data from different places.

In the context of this research, where the wearable device is expected to be on one wrist, only five datasets are valid for further work. Nevertheless, the Gravity Project dataset [25], which made use of a smartphone and an Android wearable, does not includes the TS from the on wrist sensor, only from the smartphone. Therefore, the four datasets depicted in the next item list can be used in future work for a total of 1414 TS, of which 412 include simulated falls using different sensors and sampling frequencies with different behaviour performed by up to 55 participants. Each TS will be assigned either a FALL or NOT_FALL lable, accordingly to the TS including a fall or not.

On the other hand, the presented ensembling method is one of the most simple ensembling scheme. Several improvements can be performed on this topic. Firstly, the weights can be easily fixed taking advantage of the statistical measurements of performance of the models: the better the model is, the higher the relative weight. Or the vote, because the voting scheme can also be introduced in this context.

TABLE II: Results obtained for the best model for each threshold. Different statistics are shown: the Accuracy (Acc), Kappa factor (Kp), Sensitivity (Se), Specificity (Sp), Precision or positive predictive value (Pr) and the geometric mean of Sp and Se (G). The models are feed-forward NN, rule base systems learned with C5.0 (RBS), and Support Vector Machines (SVM). ENS stands for the ensemble of the different models.

Threshold	Model	Acc	Kp	Se
2.5	NN	0.8414	0.2412	**0.8333**
	DT	**0.9288**	**0.4454**	**0.8333**
	RBS	0.8576	0.2662	**0.8333**
	SVM	**0.9288**	0.3886	0.6667
	ENS	0.8738	0.2954	**0.8333**
3.0	NN	0.8317	0.2679	**1.0000**
	DT	0.9385	0.5096	0.9167
	RBS	0.9029	0.3864	0.9167
	SVM	**0.9547**	**0.5664**	0.8333
	ENS	0.9029	0.3864	0.9167
3.09290	NN	0.8091	0.2386	**1.0000**
	DT	0.9126	0.4146	0.9167
	RBS	0.8867	0.3677	**1.0000**
	SVM	**0.9515**	**0.5484**	0.8333
	ENS	0.8997	0.3999	**1.0000**
Threshold	**Model**	**Sp**	**Pr**	**G**
2.5	NN	0.8418	0.1754	0.8375
	DT	0.9327	**0.3333**	**0.8816**
	RBS	0.8586	0.1923	0.8459
	SVM	**0.9394**	0.3077	0.7914
	ENS	0.8754	0.2128	0.8541
3.0	NN	0.8249	0.1875	**0.9082**
	DT	0.9394	0.3793	**0.9280**
	RBS	0.9024	0.2750	0.9095
	SVM	**0.9596**	**0.4545**	0.8942
	ENS	0.9024	0.2750	0.9095
3.09290	NN	0.8013	0.1690	0.8952
	DT	0.9125	0.2973	0.9146
	RBS	0.8822	0.2553	0.9392
	SVM	**0.9562**	**0.4348**	0.8927
	ENS	0.8956	0.2791	**0.9464**

Furtermore, what is really important is to introduce a new testing scenario. With all these available TS, and even considering the participant cross validation -which we do think is important-, there is enough size to have a proper training, test and validation datasets. For instance, the train dataset can be used as described in this paper to obtain the models in the feature extraction domain; the test dataset can be used in determining the best ensemble scenario, while the validation dataset would keep the same aim as in this research. It seems this experimentation might allow us to obtain a suitable FD system for this type of falls.

V. Conclusions

In this research a fall detection method has been described using a wearable device placed on a wrist measuring tri-axial accelerometry; this method is based on thresholds. The basic idea is to detect peaks of high activity; for those peaks a feature extraction stage is performed. Using publicly available datasets, a train, test and validation experimentation has been designed; the train and test stage allows to obtain models, while the validation stage mimics the performance of the system with unseen participants.

In order to reduce the number of false alarms while keeping the accuracy high, this study has proposed using ensembles. It has been found that this idea seems to perform better: not only a higher accuracy is obtained but also the number of false alarms is reduced.

Moreover, several posibilities of how to enhance these results have been discussed. On the one hand, making use of more datasets with a proper selection of data sources. On the second hand, different schemes of ensembling can be easily introduced. Besides, provided new collections of data are available, better solutions for the training, testing and validation stages have been introduced.

Finally, all these arrangements needs to be addressed, but also a careful analysis of the data gathered from the focused population is needed. As it is known, the level of activity decreases with the age, and so does the possible scenarios of a fall event. This issue must be studied using, if possible, data gathered from the target population.

References

[1] L. Rubenstein, "Falls in older people: epidemiology, risk factors and strategies for prevention," *Age Ageing*, vol. 35, pp. 37–41, 2006.

[2] F. Bagala, C. Becker, A. Cappello, L. Chiari, K. Aminian, J. M. Hausdorff, W. Zijlstra, and J. Klenk, "Evaluation of accelerometer-based fall detection algorithms on real-world falls," *PLoS ONE*, vol. 7, no. 5, p. e37062, 2012.

[3] R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," *BioMedical Engineering OnLine*, vol. 12, no. 66, 2013. [Online]. Available: http://www.biomedical-engineering-online.com/content/12/1/66

[4] S. S. Khan and JesseHoey, "Review of fall detection techniques: A data availability perspective," *Medical Engineering and Physics*, vol. 39, pp. 12–22, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1350453316302600

[5] S. Zhang, Z. Wei, J. Nie, L. Huang, S. Wang, , and Z. Li, "A review on human activity recognition using vision-based method," *Journal of Healthcare Engineering*, vol. 2017, Article ID 3090343, p. 31 pages, 2017.

[6] P. Kumari, L. Mathew, and P. Syal, "Increasing trend of wearables and multimodal interface for human activity monitoring: A review," *Biosensors and Bioelectronics*, vol. 90, no. 15, pp. 298–307, April 2017.

[7] T. Zhang, J. Wang, L. Xu, and P. Liu, "Fall detection by wearable sensor and one-class svm algorithm," in *Intelligent Computing in Signal Processing and Pattern Recognition*, ser. Lecture Notes in Control and Information Systems, I. G. Huang DS., Li K., Ed. Springer Berlin Heidelberg, 2006, vol. 345, pp. 858–863.

[8] A. Hakim, M. S. Huq, S. Shanta, and B. Ibrahim, "Smartphone based data mining for fall detection: Analysis and design," *Procedia Computer Science*, vol. 105, pp. 46–51, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050917302065

[9] F. Wu, H. Zhao, Y. Zhao, and H. Zhong, "Development of a wearable-sensor-based fall detection system," *International Journal of Telemedicine and Applications*, vol. 2016, Article ID 576364, no. 1, p. 11 pages, 2015. [Online]. Available: https://www.hindawi.com/journals/ijta/2015/576364/

[10] A. Bourke, J. O'Brien, and G. Lyons, "Evaluation of a threshold-based triaxial accelerometer fall detection algorithm," *Gait and Posture*, vol. 26, pp. 194–199, 2007.

[11] Q. T. Huynh, U. D. Nguyen, L. B. Irazabal, N. Ghassemian, and B. Q. Tran, "Optimization of an accelerometer and gyroscope-based fall detection algorithm," *Journal of Sensors*, vol. 2015, Article ID 452078, no. 452078, p. 8 pages, 2015.

[12] M. Kangas, A. Konttila, P. Lindgren, I. Winblad, and T. Jämsaä, "Comparison of low-complexity fall detection algorithms for body attached accelerometers," *Gait and Posture*, vol. 28, pp. 285–291, 2008.
[13] S. B. Khojasteh, J. R. Villar, C. Chira, V. M. González, and E. de la Cal, "Improving fall detection using an on-wrist wearable accelerometer," *Sensors*, vol. 18, no. 5, 2018.
[14] S. Abbate, M. Avvenuti, F. Bonatesta, G. Cola, P. Corsini, and AlessioVecchio, "A smartphone-based fall detection system," *Pervasive and Mobile Computing*, vol. 8, no. 6, pp. 883–899, December 2012.
[15] S. B. Khojasteh, J. R. Villar, C. Chira, V. M. González, and E. de la Cal, "Improving fall detection using an on-wrist wearable accelerometer," *Sensors*, vol. accepted for publication, pp. 1–20, 2018.
[16] A. Bourke, P. van de Ven, M. Gamble, R. O'Connor, K. Murphy, E. Bogan, E. McQuade, P. Finucane, G. Olaighin, and J. Nelson, "Evaluation of waist-mounted tri-axial accelerometer based fall-detection algorithms during scripted and continuous unscripted activities," *Journal of Biomechanics*, vol. 43, pp. 3051–3057, 2010.
[17] A. K. Bourke, J. Klenk, L. Schwickert, K. Aminian, E. A. F. Ihlen, S. Mellone, J. L. Helbostad, L. Chiari, and C. Becker, "Fall detection algorithms for real-world falls harvested from lumbar sensors in the elderly population: A machine learning approach," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 3712–3715.
[18] S. Abbate, M. Avvenuti, P. Corsini, J. Light, and A. Vecchio, *Wireless Sensor Networks: Application - Centric Design*. Intech, 2010, ch. Monitoring of human movements for fall detection and activities recognition in elderly care using wireless sensor network: a survey, p. 22.
[19] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.
[20] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "Umafall: A multisensor dataset for the research on automatic fall detection," *Procedia Computer Science*, vol. 110, no. Supplement C, pp. 32 – 39, 2017, 14th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2017) / 12th International Conference on Future Networks and Communications (FNC 2017) / Affiliated Workshops. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050917312899
[21] J. R. Villar, P. Vergara, M. Menéndez, E. de la Cal, V. M. González, and J. Sedano, "Generalized models for the classification of abnormal movements in daily life and its applicability to epilepsy convulsion recognition," *International Journal of Neural Systems*, vol. 26, no. 1650037, 2016.
[22] J. R. Villar, "Author web page, public resources." http://www.di.uniovi.es/~villar/Jose_R._Villar/Public_Data_sets/Public_Data_sets.html, 2016, last checked 16-11-2016.
[23] H. Leutheuser, D. Schuldhaus, and B. M. Eskofier, "Hierarchical, multi-sensor based classification of daily life activities: comparison with state-of-the-art algorithms using a benchmark dataset," *PLoS ONE*, vol. 8, no. 75196, 2013.
[24] E. Casilari, J.-A. Santoyo-Ramón, and J.-M. Cano-García, "Analysis of public datasets for wearable fall detection systems," *Sensors*, vol. 17, no. 1513, pp. 4324 – 4338, 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/7/1513
[25] T. Vilarinho, B. Farshchian, D. G. Bajer, O. H. Dahl, I. Egge, S. S. Hegdal, A. Lønes, J. N. Slettevold, and S. M. Weggersen, "A combined smartphone and smartwatch fall detection system," in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, Oct 2015, pp. 1443–1448.