Medical Engineering and Physics 37 (2015) 870–878

Contents lists available at ScienceDirect

Medical Engineering and Physics

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

A comparison of public datasets for acceleration-based fall detection

Raul Igual*, Carlos Medrano, Inmaculada Plaza

R&D&I EduQTech Group, Escuela Universitaria Politecnica de Teruel, University of Zaragoza, Ciudad Escolar s/n, 44003 Teruel, Spain

ARTICLE INFO

Article history:
Received 27 January 2015
Revised 19 May 2015
Accepted 24 June 2015

Keywords:
Fall detection
Accelerometers
Public datasets
Comparison
Data analysis

ABSTRACT

Falls are one of the leading causes of mortality among the older population, being the rapid detection of a fall a key factor to mitigate its main adverse health consequences. In this context, several authors have conducted studies on acceleration-based fall detection using external accelerometers or smartphones. The published detection rates are diverse, sometimes close to a perfect detector. This divergence may be explained by the difficulties in comparing different fall detection studies in a fair play since each study uses its own dataset obtained under different conditions. In this regard, several datasets have been made publicly available recently. This paper presents a comparison, to the best of our knowledge for the first time, of these public fall detection datasets in order to determine whether they have an influence on the declared performances. Using two different detection algorithms, the study shows that the performances of the fall detection techniques are affected, to a greater or lesser extent, by the specific datasets used to validate them. We have also found large differences in the generalization capability of a fall detector depending on the dataset used for training. In fact, the performance decreases dramatically when the algorithms are tested on a dataset different from the one used for training. Other characteristics of the datasets like the number of training samples also have an influence on the performance while algorithms seem less sensitive to the sampling frequency or the acceleration range.

1. Introduction

Fall incidents are a major public health problem among the older adults. Falls and the subsequent long lie period are associated with severe adverse health consequences [1–3]. The *Centers for Disease Control and Prevention* [4] quantify the direct medical cost of falls among older adults over US$30 billion per year in the United States. Every 17 s an older adult is treated in a hospital emergency department for injuries related to a fall [5]. In this context, there is a need for robust fall detectors that trigger an alert when a fall is detected [6–9].

Several techniques for fall detection have been investigated. Igual et al. [6] classified the existing fall detection studies into two categories: context-aware systems [10] and acceleration-based wearable devices [11]. One of the characteristics of acceleration-based studies is that they report high detection rates. For example, sensitivity and specificity are reported respectively as 97.5% and 100% by Kangas et al. [12], 94.6% and 100% by Bourke et al. [13], 98.6% and 99.6% by Yuwono et al. [14]; and 100% and 100% by Abbate et al. [15]. Other fall detection studies provide similar performances [16,17]. It should be noted that the detection rates provided by all these studies are very high. However, many authors on this field have noticed strong difficulties when comparing different acceleration-based studies [6,18]. This is due to the fact that each study uses its own dataset composed of simulated falls and ADL. Therefore, it is not clear whether the declared results are influenced by the specific dataset used and it is not possible to perform a fair comparison since the datasets used to provide a measure of the detection performances are different in each study.

In this regard, several authors have identified the need for having public datasets [19,20]. Some efforts have been performed in this direction since several datasets were made publicly available in the recent years: *DLR* [21] published in 2011, *MobiFall* [22] available in 2013 and *tFall* [20] uploaded in 2014 (the study of Fudickar et al. [23] cites another public dataset but it seems that it cannot be downloaded currently). Although these three datasets can be freely accessed, there is no study focused on comparing them. Therefore, some important questions remain unanswered: Can the public datasets be used indistinctly?; Are there any differences among them?; Is the performance of the fall detection algorithms affected by the specific selected dataset?

In this regard, the general goal of this paper is to compare in a fair way the existing public datasets (Fig. 1). For that purpose, the following specific objectives are stated:

(1) To check whether or not the performance of a given algorithm depends on the selected dataset;

* Corresponding author. Tel.: +34 978 64 53 62; fax: +34 978 61 81 04.
E-mail addresses: rigual@unizar.es, rauligual@gmail.com (R. Igual), ctmedra@unizar.es (C. Medrano), inmap@unizar.es (I. Plaza).

http://dx.doi.org/10.1016/j.medengphy.2015.06.009
1350-4533/

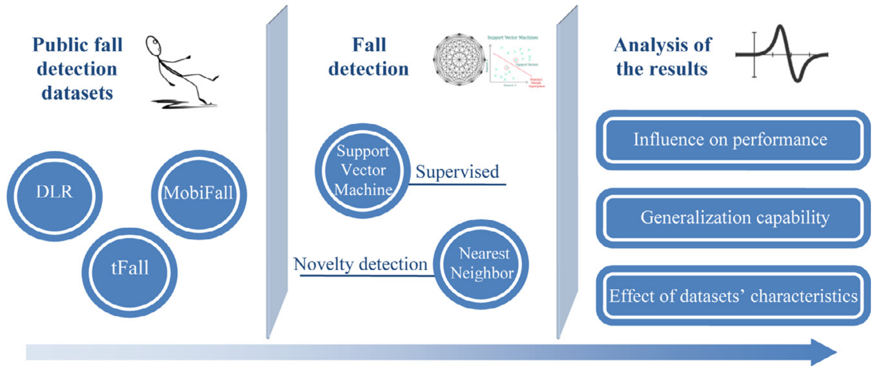

Fig. 1. General schema of the study.

Table 1
Features of the public fall detection datasets.

		DLR	MobiFall	tFall
Experiments	**No. subjects**	16	11	10
	Device	Xsens MTx	Samsung Galaxy S3	Samsung Galaxy Mini
	Position	Belt	Pocket	Pocket
	Types of falls	Not specified	Forward-lying, front-knees-lying, sideward-lying and back-sitting-lying	Forward, forward straight, backward, lateral left and right, sitting on empty air, syncope and forward fall with obstacle
	Types of ADL	Sitting, standing, walking, running, jumping and lying	Standing, walking, jogging, jumping, stairs up, stairs down, sitting on a chair, step in a car and step out a car	Real-life activities
Samples	**No. ADL**	1077	831	7816
	No. falls	53	132	503
	Sampling frequency	100 Hz	100 Hz	50 Hz
	Acc. range	7 g	2 g	2 g

(2) To compare the generalization capability of the public datasets. Generalization capability refers to the ability of a system trained under some conditions to work under different conditions;
(3) To determine whether some of the datasets' parameters affect the performance of the fall detection algorithms.

2. Materials and methods

2.1. Datasets

As a result of an extensive literature search, we identified three public datasets presenting acceleration samples of falls and ADL: *DLR* [24], *MobiFall* [25] and *tFall* [26]. These three datasets were collected by different research institutions, each conducting the experiments in a particular fashion. These datasets were selected since, to the best of our knowledge, they are the only ones that are publicly available to the scientific community.

2.1.1. DLR dataset

This dataset was made publicly available by the *Institute of Communications and Navigation* of the *German Aerospace Center* (*DLR*). The dataset was collected from 16 male and female subjects aged between 23 and 50 and annotated manually by an observer. In total it contains about 4.5 h of labeled falls and activities (Table 1). Each participant performed a different number of ADL and falls. To capture the motion data, the *Xsens MTx* inertial measurement unit with a single tracker placed on the belt was used. The data were sampled at 100 Hz and the measurement unit had an acceleration range of at least 7 g.

2.1.2. MobiFall dataset

This dataset was developed by the *Biomedical Informatics & eHealth Laboratory* of the *Technological Educational Institute of Crete*. The *MobiFall* dataset contains data from 11 volunteers: six males and five females (age range: 22–36). Nine participants performed falls and ADLs, while two performed only the falls. On the one hand, each participant performed four types of falls which were repeated three times per subject. On the other hand, nine types of ADL were simulated (Table 1). Specifically, a Samsung Galaxy S3 device with the *LSM330DLC* inertial module (3D accelerometer and gyroscope) was used to capture the motion data. The device was located in a trouser pocket freely chosen by the subject in any random orientation. The range of the accelerometer was 2 g and the data were acquired at 100 Hz.

2.1.3. tFall dataset

This dataset was developed by the *EduQTech (Education, Quality and Technology)* group of the *University of Zaragoza*. Ten people were involved in the data collection process (seven males and three females, whose ages ranged from 20 to 42). The simulation set consisted of eight different types of falls (Table 1). Each fall was repeated three times per subject. The ADL collection process was carried out under real-life conditions. ADL were recorded in the subjects' real world environment while they performed their daily lives. Each subject was monitored during at least one week. Only ADL over a given threshold (1.5 g) were recorded. At the end of the experience, an average number of about 800 records per subject (6 s length) were obtained. The data were acquired using Samsung Galaxy Mini phones at 50 Hz and with a range of 2 g. In the fall study, participants carried a phone in both their two pockets.

2.2. Fall detection algorithms used to compare the datasets

It is clear that the comparison can depend on the algorithm. Therefore, we have selected two algorithms representing different approaches to fall detection.