Open Lab Nursing Activity Recognition Challenge

Paula Lago[1,a) Shingo Takeda [1,b) Alia Sayeda Shamma[1,c) Tittaya Mairittha [1,d)
Nattaya Mairittha [1,e) Sozo Inoue[1,2,f)

Abstract: Although activity recognition has been studied for a long time now, research and applications have been mainly focused on physical activity recognition. Even if many application domains require the recognition of more complex activities, for example daily living activities or work-related activities, research on such activities has attracted less attention. One reason for the gap in research for complex activities is the lack of datasets to evaluate and compare different methods. To promote research in such scenarios, we will organize the *Open Lab Nursing Activity Recognition Challenge* focusing on the recognition of complex activities related to the nursing domain. Nursing domain is one of the domains that can benefit enormously from activity recognition but has not been researched due to lack of datasets. The competition will use the *Open Lab Nursing Activities Dataset*, featuring 7 activities performed by 8 subjects in a controlled environment with accelerometer sensors, motion capture and in-door location sensor. The dataset was collected at the Smart Life Care Society Creation Unit in Kyutech, Japan, based on the collaboration between Kyutech and Carecom. Co., Ltd. In this paper, we describe the data collection experiments and the dataset.

Keywords: activity recognition, nurse care, machine learning, data collection

1. Introduction

Although activity recognition has been studied for a long time now, research and applications have been mainly focused on physical activity recognition [1, 2, 6]. Many commercial products can now recognize when we walk, cycle or run, count steps and give a general overview of how active we have been on a day. However, research in applications to domains like healthcare, although advertised and highly regarded as important [4], is not very mature. One reason is that the activities are more complex and hard to analyze. Research in these areas is even more difficult due to the lack of publicly data that can be used by researchers to propose and compare the performance of different methods.

In this challenge, we aim at bridging the gap for nurse care activity recognition. Complex activity recognition in health care applications has been focusing on recognition of patient activities [3, 5, 7] and overlooked nurses and caregivers. Recognition of nurses activities can have many applications like automatic record creation to reduce documentation time, checking compliance with care routines for a given patient and identification of risk activities that require special, for example hand washing after taking blood samples.

Nurse care activity recognition is challenging because, unlike other activity settings in which the user is doing an activity, nurses usually perform some activity *TO* a patient. For example, they give a drink to the patient instead of just drinking. This feature introduces challenges that are not studied in other activity recognition applications. For example, the activities can be performed differently by the same nurse depending on the patient receiving the care. In this case, intra-class variability depends not only on the subject, as in other domains, but also on the receiving patient.

To study the feasibility and challenges of implementing activity recognition using movement and position for this domain, we have performed an experiment to measure nurse activities in a controlled environment. The experiment was done in collaboration with CARECOM, a leading company of the Nursing Support System Industry with 60 years of history and experiences in Japan. The controlled environment allows to eliminate challenges related to sensing, such as sensor position and orientation, and enables the research to focus on the challenges related to the types of activities.

In this experiment, we collected a multimodal dataset of 6 activities with 8 subjects, 5 repetitions each, yielding about 240 activity sequences and 407 recorded minutes. We recorded each activity using accelerometers, meditag data (bluetooth based localization and pressure sensor), and motion capture data including 25 body markers.

In this paper, we present the collection experiment details, a summary of the collected data and we describe the activity recognition challenge which will be hosted using the collected dataset. We first describe the experiment settings (Section 2) and then we summarize the collected data (Section 3). Next, we introduce the activity recognition challenge (Section 4) using the collected data to motivate researchers in this domain. Finally, we conclude this work in Section 5

1 Kyushu Institute of Technology, Kitakyushu, Fukuoka, Japan
2 Riken AIP
a) paula@mns.kyutech.ac.jp
b) takeda@sozolab.jp
c) alia@sozolab.jp
d) fon@sozolab.jp
e) fah@sozolab.jp
f) sozo@brain.kyutech.ac.jp

2. Experiment Description

In this section, we describe the nursing activities selected for the experiment, describe the experimental settings under which the data was collected and summarize the dataset.

2.1 Nursing activities selection

The dataset contains data about 6 nursing activities, related to hospital patient care. The choice of activities was done by CARECOM. In Table1, we describe every activity.

2.2 Experimental settings

The dataset was recorded in the Smart Life Care Unit of the Kyushu Institute of Technology in Japan (Figure 1). This unit is located in the Wakamatsu Campus of the University and has motion capture equipment, nursing equipments, as well as other sensors like EMG, EEG, eye movement sensors and others.

Fig. 1: Smart Life Care Unit of Kyushu Institute of Technology

Data collection was done in three days total. Each day, two or three professional nurses came to the laboratory and were asked to perform the six activities of Table 1. All nurses agreed to the sensor data collection for this experiment.

For each activity, instructions about the procedure were given, so that the activity comprises all necessary steps. One person acted as patient so the activity is performed in naturalistic way. Each nurse performed the activity in her/his preferred way, pace and order, where applicable.

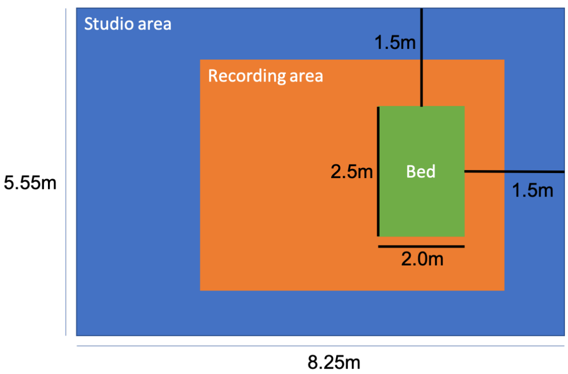


Fig. 2: Recording Studio area plan

The room was equipped with hospital bed, desk, and wheeled cart for the nurse to take the necessary equipment. All instruments like drip, gauze, diaper, etc. were provided and used by the nurses during data collection. A plan of the studio area where the recordings were done is shown in Figure 2 with corresponding measurements.

2.3 Sensors used

We collected data from motion capture, meditag and accelerometer sensors. Video was also recorded for each run, but it will not be released due to privacy concerns of the participants. The description of each data source is given below.

Motion Capture: We use the motion capture system form Motion Analysis Company[*1]. The setup used consists of 29 body markers located as in Figure 4. The markers are tracked using 16 infrared cameras (Kestrel Digital Real Time as in Figure 3). The three dimensional position of each marker is recorded with a frequency of 100Hz. The markers may be labeled incorrectly in some cases due to the complex setting.

Fig. 3: Motion capture camera

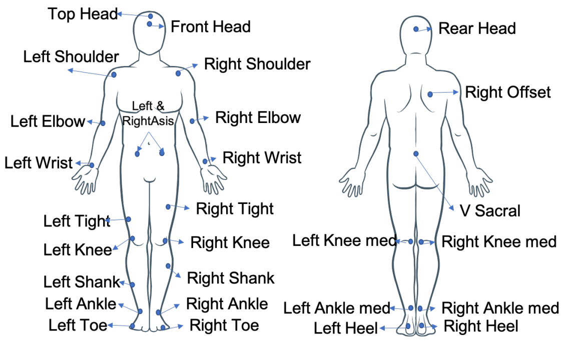


Fig. 4: Motion capture markers

Meditag Sensor: The meditag is a bluetooth based in-door localization sensor. Four receivers were installed in the room and

*1 http://motionanalysis.com/movement-analysis/

Table 1: Description of the activities

Id	Activity Name	Activity Purpose
2	Vital signs measurements	Confirm the signs of human life by measuring respiration, pulse, temperature and blood pressure. Consciousness may be included by asking questions.
3	Blood Collection	Collect blood from the body to know the progress of diagnosis and treatment of disease.
4	Blood Glucose Measurement	Measure blood sugar to control it. This procedure is indispensable for diseases that require sugar control such as diabetes.
6	Indwelling drip retention and connection	The purpose of indwelling placement is: 1. Refill the body with insufficient fluid. 2. Correct and maintain electrolyte balance. 3. Prosperity for patients who can not take orally. Provide support. 4. Administer the necessary medications.Etc.
9	Oral care	Keep the mouth clean to keep the health of the whole body as well as in the mouth. Other purposes are: 1. Promote the secretion of saliva. 2. Prevent infection and fever. 3. Prevent dementia. 4. Prevent aspiration pneumonia. 5. Prevent deterioration of oral function
12	Diaper exchange and cleaning of area	Wash to keep the genital area clean for patients who can not take a bath or for patients wearing diapers. If a bladder catheter is indwelling, it is done as prevention of retrograde urinary tract infection.

the nurse carried a bluetooth beacon on her right chest pocket. It measures two dimensional position (x and y) of the beacon in meters and air pressure in mHg with a sampling rate of approximately 20 Hz.

Accelerometer sensor: We used a Freetel Priori 3 smartphone carried in the right chest pocket of the nurse in upright position. The sampling rate is approximately 4 Hz. This sensor measures three dimensional inertial acceleration in m/s^2.

3. Data Description

In this section, we summarize the data collected as a result of the experiment. In total, we recorded 240 activity sequences and 407 minutes. The distribution of recorded minutes for each activity is shown in Figure 5. Even if we collected 40 samples (5 per user) of each activity, the duration of the activities is not the same, so the time distribution is not equal. We show the distribution of the duration of each activity in Figure 6. From the figure, we can see that the duration of all activities is between 1 and 4 minutes in average. Notice that the activities with longer duration (Activity 3 and 12) also have the highest number of recorded minutes.

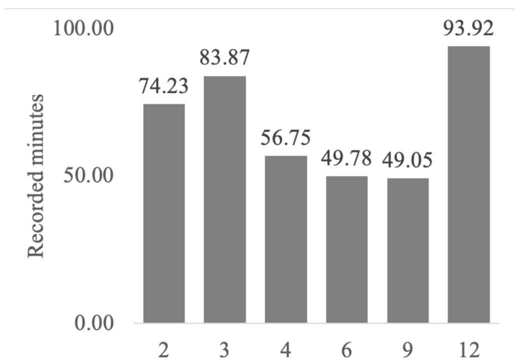


Fig. 5: Recorded minutes of each activity. Different activity duration raise an imbalance in time distribution even if the number of samples for each activity is the same.

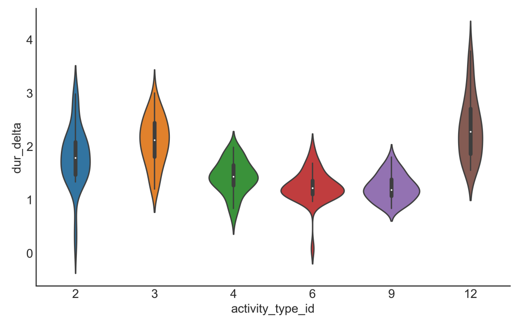


Fig. 6: Distribution of the duration of each activity class.

However, the recorded minutes for each nurse is almost equally distributed (Figure 7). The difference in recording times is due to variations in activity performance, some nurses taking more time than others, and not due to differences in number of samples per activity. The number of observations by each sensor is shown in Figure 8. Due to the big sampling rate differences, there are more measurements for the motion capture than for other sensors. As examples of the data collected, in Figure 11 we show one example of the accelerometer data collected for each activity. Similarly, we show one example for each activity of the position

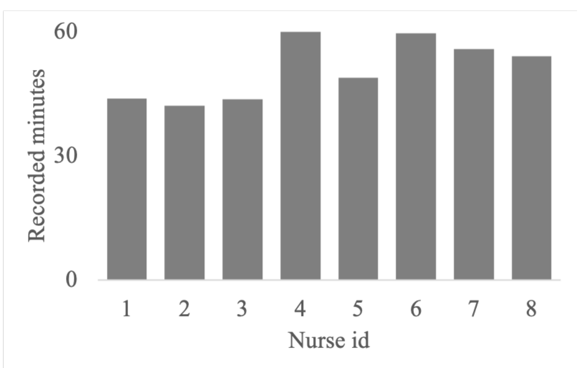


Fig. 7: Recorded minutes by user.

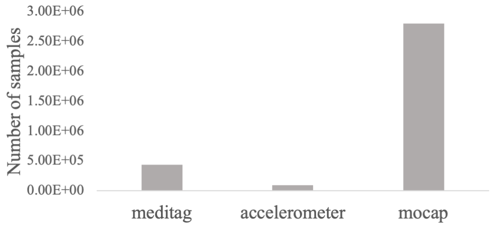


Fig. 8: Number of observations by each sensor source.

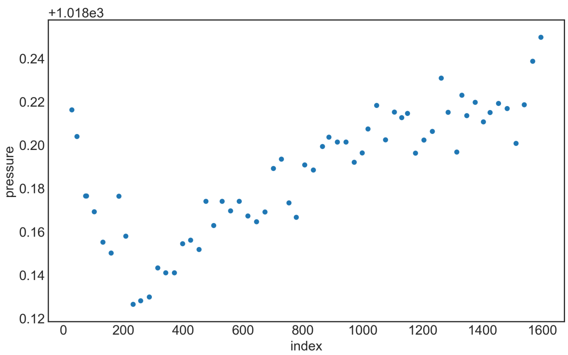


Fig. 9: Pressure measured by meditag sensor

data collected by the meditag sensor in Figure 13. Example data from the pressure sensor is shown in Figure 9.

4. Challenge description

We will host an Activity Recognition Challenge using the collected dataset. The goal of the challenge is to recognize with high accuracy the activity being performed by a nurse in a given minute.

For this challenge, each activity recording has been segmented into 1-minute segments. The goal of the challenge is to classify each segment's activity. Segments have been given random identifiers, so their sequence is unknown. When windowing, at the end of an activity the data may not be enough for a complete 1-minute. For this reason, we have removed segments with duration less than 20 seconds, but some segments may be shorter than 1-minute.

The dataset has been divided into two parts: one labeled dataset which contains data from 6 users and one unlabeled dataset containing data from the remaining 2 users.

4.1 Data organization

The dataset is organized in two folders: train and test. Train folder contains labeled data used for training and the test folder contains the data that participants should estimate. Root folder contains both folders and a README file. The train and test folders are structured as illustrated in Figure 14.

```
root
└─train
   └─accelerometer.csv
   └─meditag.csv
   └─mocap
      └─segment1.csv
      └─segment2.csv
      └─other files
   └─activities.csv
└─test
   └─accelerometer.csv
   └─meditag.csv
   └─mocap
      └─segment300.csv
      └─segment301.csv
      └─other files
└─README
```

Fig. 14: Dataset file structure

The accelerometer.csv contains all the accelerometer sensor readings. An example of the file structure is illustrated in Figure 15. Each line of this file has a segment id indicating the identifier of each 1-minute segment, a sequence id indicating the sequential order of the measurements in the segment and the x, y and z acceleration as measured by the smartphone.

```
segment_id,seq_id,x,y,z
13,1,-1.838,-9.5,0.023
13,2,-1.838,-9.5,0.023
13,3,-1.11,-9.577,0.112
13,4,-1.11,-9.577,0.112
13,5,-1.11,-9.577,0.112
13,6,-1.11,-9.577,0.112
13,7,-1.11,-9.577,0.112
```

Fig. 15: Accelerometer file example

The meditag.csv contains the readings of the meditag sensor as illustrated in Figure 16. As in the accelerometer file, each line has a segment id, a sequence id and the x, y and pressure readings.

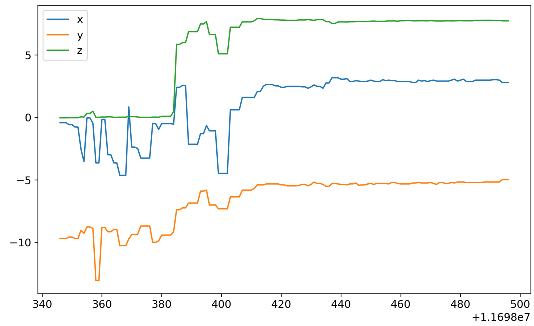
```
segment_id,seq_id,x,y,pressure
13,1,-1.04,-3.97,1018,177
13,2,-1.04,-3.97,nan
13,3,-1.04,-3.97,nan
13,4,-1.04,-3.97,nan
13,5,-1.04,-3.97,nan
13,6,-1.04,-3.97,nan
13,7,-1.04,-3.97,nan
```

Fig. 16: Meditag sensor file example

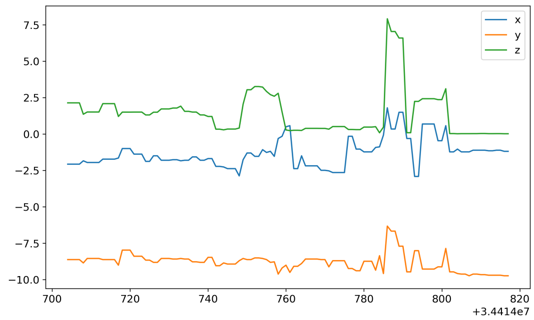
The motion capture files are organized in a folder due to their size. The folder contains one file per segment named segment_[s_id].csv. Each file contains the segment id, the sequence number of the reading and the position (X,Y,Z) of each marker, as illustrated in Figure 17. As in the accelerometer file, each line has a segment id, a sequence id and the x, y and pressure readings. The segment identifiers correspond in all files.

```
seg_id,seq_id,m1_x,m1_y,m1_z, m2_x ...
13,1,-17.29,852.04,1665.84,49.69...
13,2,-17.29,852.03,1665.82,49.72...
13,3,-16.53,852.58,1665.42,49.16...
13,4,-16.50,852.57,1665.41,49.37...
13,5,-16.50,852.57,1665.42,49.39...
13,6,-16.45,852.54,1665.44,49.45...
```

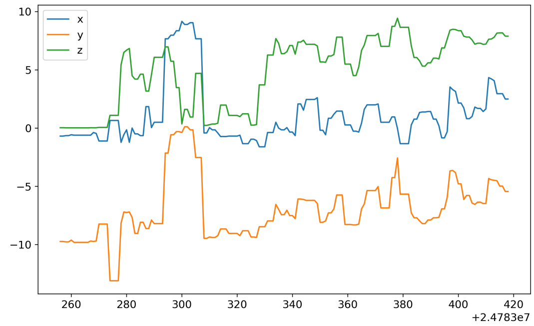
Fig. 17: Mocap sensor file example



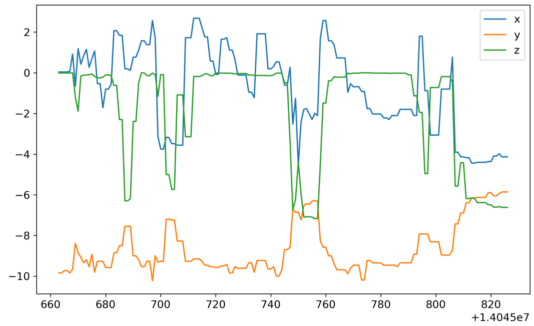
Example of accelerometer data for Vitals signs measurement (Activity 2)



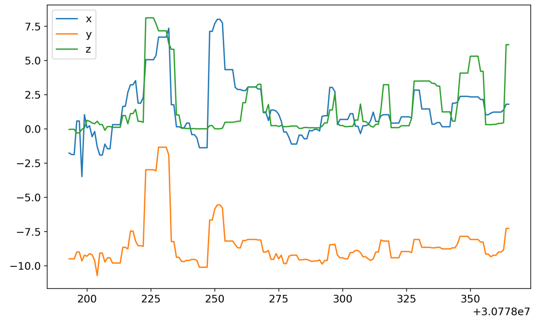
Example of accelerometer data for Blood Collection Activity (Activity 3)



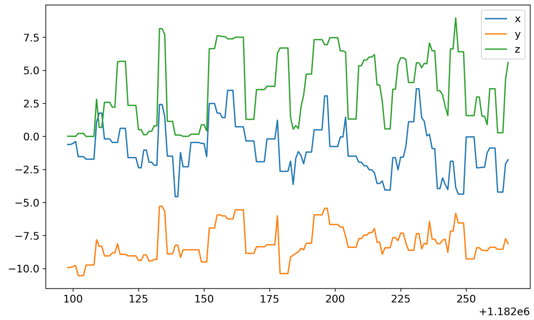
Example of accelerometer data for Blood Glucose Measurement Activity (Activity 4)



Example of accelerometer data for Indwelling drip retention and connection (Activity 6)



Example of accelerometer data for Oral Care Activity (Activity 9)



Example of accelerometer data for Diaper Exchange Activity (Activity 12)

Fig. 11: Accelerometer data examples for each activity

The ACTIVITIES.csv file contains the activity id for each segment as illustrated in Figure 18. Segment identifiers correspond in all files, so that all sensors sources can be combined in a feature vector.

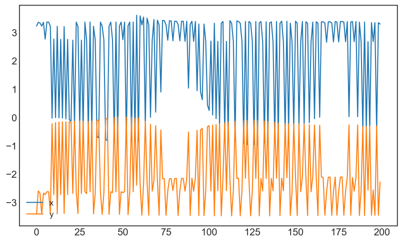
```
segment_id,activity_id
1,2
2,3
3,12
4,4
5,2
6,6
7,9
```

Fig. 18: Activities file example

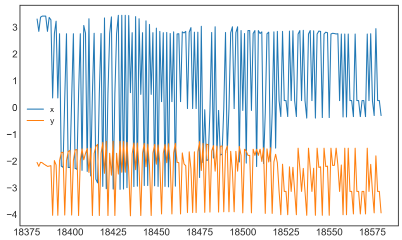
All data sources are available for the challenge. Data is provided as is, with no pre-processing from our part other than the windowing into 1-minute. Therefore, data can be noisy in some ways. In particular, due to the different sampling rates, there are many missing values for the pressure measurement of the meditag sensor. Participants can use one source, several sources or all data sources.They are encouraged to use different pre-processing pipelines. Since the position data is dependent on the environment, participants are encouraged to design features that can be re-used in a different setting with good accuracy.

4.2 Challenge participation

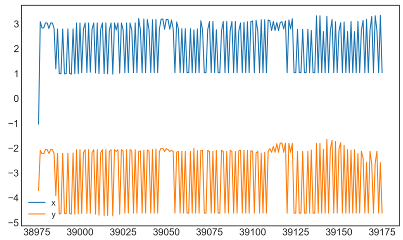
The challenge will be open until June 30, 2019. Participants should follow the instructions in the webpage https://hasc-nurse-challenge.github.io/. The results will be evaluated based on accuracy metric. Participants should submit a paper to the HASCA workshop, to be held during the Ubicomp Conference, describing the data sources selected, their preprocessing pipeline, the features used and their results. All submissions must follow the HASCA format (up to 10 pages).



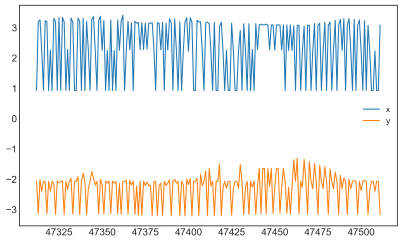
Example of meditag data for Vitals signs measurement (Activity 2)



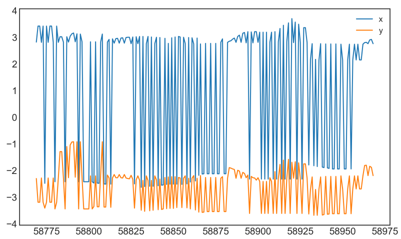
Example of meditag data for Blood Collection Activity (Activity 3)



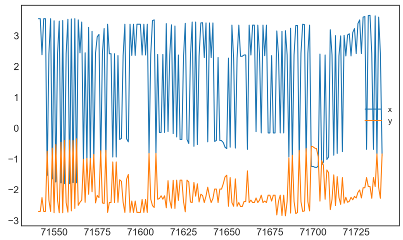
Example of meditag data for Blood Glucose Measurement Activity (Activity 4)



Example of meditag data for Indwelling drip retention and connection (Activity 6)



Example of meditag data for Oral Care Activity (Activity 9)



Example of meditag data for Diaper Exchange Activity (Activity 12)

Fig. 13: Meditag indoor location data examples for each activity

5. Conclusion

In this paper, we have presented an experiment and dataset for nurse activity sensing to bridge the gap on activity recognition research for complex activities. The collected dataset features rich, multimodal sensor data for a set of common nurse activities. The goal of our experiment is to evaluate the feasibility of nurse activity recognition based on body movement and location with respect to a hospital room. The data was collected in a controlled environment, so the focus is on the activities and not on the challenges related to sensing, such as device orientation/position changes.

We open the data through the Open Lab Nurse Activity Recognition Challenge, hosted as part of the 7th HASCA Workshop to be held during Ubicomp 2019. The challenge participants can use the data freely and are encouraged to submit papers to this workshop.

The application of activity recognition technology in the healthcare domain opens innovative services for improving quality of care, work satisfaction, and others.

References

[1] Stefano Abbate, Marco Avvenuti, Francesco Bonatesta, Guglielmo Cola, Paolo Corsini, and Alessio Vecchio. A smartphone-based fall detection system. *Pervasive and Mobile Computing*, 8(6):883 – 899, 2012. Special Issue on Pervasive Healthcare.

[2] Ling Bao and Stephen S. Intille. Activity Recognition from User-Annotated Acceleration Data. In A. Ferscha and F. Mattern, editors, *Pervasive Computing. Pervasive 2004. Lecture Notes in Computer Science, vol 3001.*, Berlin, Heidelberg, 2004. Springer, Berlin, Heidelberg.

[3] Saisakul Chernbumroong, Shuang Cang, Anthony Atkins, and Hongnian Yu. Elderly activities recognition and classification for applications in assisted living. *Expert Systems with Applications*, 40(5):1662–1674, apr 2013.

[4] J. Favela. Behavior-aware computing: Applications and challenges. *IEEE Pervasive Computing*, 12(3):14–17, July 2013.

[5] Paula Lago, Fréderic Lang, Claudia Roncancio, Claudia Jiménez-Guarín, Radu Mateescu, and Nicolas Bonnefond. The contextact@a4h real-life dataset of daily-living activities. In Patrick Brézillon, Roy Turner, and Carlo Penco, editors, *Modeling and Using Context*, pages 175–188, Cham, 2017. Springer International Publishing.

[6] Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. A practical approach to recognizing physical activities. In Kenneth P. Fishkin, Bernt Schiele, Paddy Nixon, and Aaron Quigley, editors, *Pervasive Computing*, pages 1–16, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[7] M. Stikic, D. Larlus, S. Ebert, and B. Schiele. Weakly supervised recognition of daily life activities with wearable sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2521–2537, Dec 2011.

