

Personalized Posture and Fall Classification with Shallow Gated Recurrent Units

Paul Compagnon, Grégoire Lefebvre, Stefan Duffner, Christophe Garcia

► **To cite this version:**

Paul Compagnon, Grégoire Lefebvre, Stefan Duffner, Christophe Garcia. Personalized Posture and Fall Classification with Shallow Gated Recurrent Units. 32nd IEEE CBMS International Symposium on Computer-Based Medical Systems, Jun 2019, Cordoue, Spain. hal-02095355

HAL Id: hal-02095355

<https://hal.archives-ouvertes.fr/hal-02095355>

Submitted on 10 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Personalized Posture and Fall Classification with Shallow Gated Recurrent Units

Paul Compagnon^{*†}, Grégoire Lefebvre^{*}, Stefan Duffner[†] and Christophe Garcia[†]

^{*}Orange Labs, Grenoble, France

{paul.compagnon, gregoire.lefebvre}@orange.com

[†]LIRIS, UMR 5205 CNRS, INSA Lyon, France

{stefan.duffner, christophe.garcia}@liris.cnrs.fr

Abstract—Activities of Daily Living (ADL) classification is a key part of assisted living systems as it can be used to assess a person autonomy. We present in this paper an activity classification pipeline using Gated Recurrent Units (GRU) and inertial sequences. We aim to take advantage of the feature extraction properties of neural networks to free ourselves from defining rules or manually choosing features. We also investigate the advantages of resampling input sequences and personalizing GRU models to improve the performances. We evaluate our models on two datasets: a dataset containing five common postures: sitting, lying, standing, walking and transfer and a dataset named MobiAct V2 providing ADL and falls. Results show that the proposed approach could benefit eHealth services and particularly activity monitoring.

Keywords-Gated Recurrent Units, Posture Detection, Fall Detection, eHealth, Inertial Measurement Unit

I. INTRODUCTION

As life expectancy increases, more and more elderly people show difficulties in their every day life, and allowing them to stay at home is a social and public health issue¹. Many people even struggle performing basic need activities. They are also particularly exposed to chronic diseases: diabetes, cancer, psychological and cognitive disorders, heart diseases, Parkinson and Alzheimer, *etc.* They are finally vulnerable in simple daily life activities where they could fall, make a wrong move or lose attention. Furthermore, 15% of the world population lives with a form of handicap, between 2 and 4% live with severe disabilities². All these persons could benefit from eHealth services and particularly activity monitoring [1]. Actigraphy is the process of recording the every day life activities of a subject using sensors (inertial sensors, for instance). Instead of organizing regular visits at the hospital, the patient can be monitored in his/her house with several upsides: it improves the quality of life of the patient and shorten hospital stays while facilitating the diagnosis as important data are collected in the usual environment of the patient. Off course, clinical visits are indispensable but can only take a snapshot of the

patient’s condition and may occur too late during the disease development [2].

Inertial data are particularly interesting as they are nowadays easily available from smartphone sensors (or smart watches, *etc.*) which can be carried without stigmatizing the person and are judged less intrusive at home. These data can then be used to provide eHealth services regarding the recorded level of activity and the distribution of those activities during the day. Traditionnaly, the level of autonomy has been evaluated thanks to several criteria related to the Activities of Daily Living (ADL). This evaluation appears as a pertinent factor for the clinical evaluation of elderly people [3]. Nowadays, it is possible to observe, with the help of actigraphy systems, changes in a person behavior and so the possible loss of autonomy. When processed through machine learning algorithms, data obtained with actigraphy systems allow to perform activity classification and prediction and to automatically detect changes. To be able to eventually equip people with such systems, it requires high prediction accuracy, particularly on critical events such as falls.

In this paper, we improve ADL and fall classification from inertial data sequences. Our aim is to use as little as possible outside expertise because we design our system to easily adapt to a new user or a new activity. Thus, we prefer not to define expert rules and not to extract explicit features from input sequences in order to benefit as much as possible from the model’s capability to extract and learn automatically saillant features. We then propose personalized shallow Gated Recurrent Unit (GRU) [4] models as they hold the right properties to train and embed the samples of each individual (*i.e.* close to real actigraphy systems). We will show that resampling sequences during the training phase can allow to better recognize classes, notably when a category is represented by short and rare sequences (*e.g.* fall detection).

The paper is organized as follows. We give in section II an overview of the human activity recognition by focusing on inertial data and machine learning. We then introduce GRU networks in section III-A and explain inertial sequential data specificities in section III-B. The section IV is dedicated to the description of the experimental datasets and setups and

¹www.who.int/ageing/publications/world-report-2015

²www.who.int/disabilities/world_report/2011/report

exposes our results. Finally, conclusions and perspectives are drawn in Section V.

II. RELATED WORK

Human Activity Recognition is a very broad computer science field which aims to recognize what a person is doing by exploiting data. It has numerous applications: from crime detection on video surveillance images to gesture recognition when performing a physical activity. It can be performed in several contexts (Lara *et al.* [5] listed seven types: ambulation, transportation, phone usage, exercise/fitness, military, upper body), using different types of data (*e.g.* photos, videos, sounds but also smart home data, smart phone communications data or inertial data) and with different approaches (*i.e.* time series analyzing, rule-based models, statistical models, machine learning and especially neural networks). We focus our state of the art on papers performing ADL recognition on inertial data with machine learning algorithms. The approaches can be divided into two categories: those using handcrafted features from the data such mean or maximum amplitude, energy, entropy, *etc.* and those automatically extracting discriminant features.

Firstly, we mention papers exploiting feature vectors from the signal. Bao *et al.* [6] used a decision tree to classify 20 activities including daily household activities with data coming from five accelerometers placed at different body locations. They classified vectors composed of several extracted features such as mean, energy, *etc.* They observed that user-specific models could lead to better results.

Another paper by Yang *et al.* [7] presents a system which can exploit the data of a triaxial accelerometer to recognize eight common daily activities. They divide the classification into two subtasks: separate dynamic and static activities with a preclassifier using two features: average energy and signal magnitude area. The system then performs activity classification using a feed forward neural network trained on feature vectors extracted from the signal. The system has the drawback of comprising many parameters and of using a cascade of classifications. This requires very good performances at each step in order to not affect the overall performances.

Likewise, Jafari *et al.* [8] conceived a system to distinguish the fall of a human from its normal behavior with data coming from a three-axial accelerometer. The data first pass through several preprocessing operations, then some possibilities are eliminated with statistical rules (static, walking, *etc.*). Finally, 4 features are extracted (*e.g.* postural orientation, singular value decomposition, skewness and maximum amplitude change) from the signal and classified into 4 movement transitions with a neural network. The transition from standing to lying corresponding to a fall. The neural network is only used for classification whereas

it can be trained to also extract the pertinent features for classification.

Chatzaki *et al.* [9] proposed a dataset called MobiAct V2 of 12 ADL and 4 falls. They performed activity and fall recognition on it with the goal to develop the most effective pipeline in terms of accuracy. To do so, the authors conducted an exhaustive study to find the best features (for example, they found that the spectral centroid was quite essential). They achieved their best accuracy score using decision trees and k-nearest neighbors. We remark that the accuracy was lower for fall categories.

Secondly, we present some papers employing automatic feature extraction algorithms. Zeng *et al.* [10] used a Convolutional Neural Network to automatically extract features from sequences of accelerometer data coming from a mobile phone and performed activity classification using k-nearest neighbors. Their approach outperformed common feature extraction approaches such as principal component analysis.

In another work, Lefebvre *et al.* [11] used Bidirectional Long-Short Term Memory (LSTM) to recognize symbolic gestures from raw inertial data (accelerometer and gyroscope). They compared their approach to a Hidden Markov Model (HMM) and Dynamic Time Warping on 3 experimental paradigms including personalized models. They achieved better performances on two of those but not on personalized models on which HMM performed best.

Likewise, Di Pietro *et al.* [12] proposed a new neural network model called MIXed hiSTory Recurrent Neural Networks (MIST) which they tested on MobiAct V2, among others. They did not select features and learn directly from the raw data. They separated users into fixed train, validation and test groups and repeated the same experiment 50 times and kept the 5 best results. Their approach both requires less computation and produce better results than LSTM.

All the previous works tested their approach on different datasets and the lack of reference datasets is a known issue among the activity recognition community. It is due to the variety of contexts in which it can be performed and also the variety of data that can be used. We chose to concentrate on two datasets on which results have already been published so we can make comparisons. We trained our GRU models on a dataset called Postures used by the authors of [13]. Unlike this article, we trained personalized neural network models and so did not define expert rules to classify the postures. The second dataset is MobiAct V2 [9]. Our approach is different from the path followed in the original article as we do not select features and learn directly from the raw resampled data with neural network approaches instead of decision trees. In [12], the authors trained also RNN architectures on the raw sequences of MobiAct v2. Our approach differentiates itself as we used shallow GRU

(two or three layers) instead of deep LSTM/ MIST RNN and we propose a resampled version of the dataset to predict not only ADL but exclusively falls.

III. RECURRENT MODELS FOR ADL CLASSIFICATION

In the following, we will present our ADL classification pipeline which does not require to extract explicit features from the signal or to define expert rules. For those reasons, we chose to investigate GRU RNN.

A. Gated Recurrent Units

Recurrent Neural Networks are similar to feedforward networks but possess recurrent connections which give the ability to map an input sequence to an output sequence while at each step taking the information of previous steps into account. This enables the network to not only extract inter-signal correlations but also intra-signal ones and thus to detect more complex patterns. Let x be the sequence of inputs of the network and x_i be the input at time i . A recurrent network computes \hat{y} , the sequence of outputs following this equation:

$$\hat{y}_t = f(W_x x_t + W_h h_{t-1} + b_h) \quad (1)$$

Where W_x and W_h designate synaptic weights matrices, b the bias and h the recurrent connection, the hidden state coming from the previous step. f is an activation function, generally \tanh . It is a well-known issue that these networks struggle with long-term dependencies, that is to say when they need to learn to retain information during a long time [14]. This is the so-called vanishing gradient problem which has been partly solved by LSTM [15]. GRU [4] can be viewed as a simplified version of the classic LSTM approach, while showing similar performances on some classical recurrent models tasks [16], namely speech and music modeling. A GRU will modify the way the vector h_t is computed (*i.e.* it modifies the hidden layer):

$$h_t = (1 - z_t) \tilde{h}_t + z_t h_{t-1} \quad (2)$$

Similarly to the cell of the LSTM, h_t is updated by forgetting old content and directly adding some new. z_t is called the update gate and is computed according to the following equation:

$$z_t = \sigma(W_{iz} x_t + b_{iz} + W_{hz} h_{t-1} + b_{hz}) \quad (3)$$

where σ is the logistic sigmoid function. We have the following relation for \tilde{h}_t named the new gate:

$$\tilde{h}_t = \tanh(W_{i\tilde{h}} x_t + b_{i\tilde{h}} + r_t (W_{h\tilde{h}} h_{t-1} + b_{h\tilde{h}})) \quad (4)$$

where finally r_t is the reset gate computed similarly as z_t :

$$r_t = \sigma(W_{ir} x_t + b_{ir} + W_{hr} h_{t-1} + b_{hr}) \quad (5)$$

To perform classification, this model is trained with the negative log likelihood loss. The size of the output vector is equal to the number of classes to predict and this loss will make the output a probability distribution over the different classes (softmax). RNN are trained using an adapted version of the back propagation called the Back Propagation Through Time (BPTT). In fact, when a RNN is unrolled, it behaves nearly as a feedforward network and the BPTT just takes into account the activations of the network for each time step.

B. Making Predictions from Sequential Inertial Data

Postures, ADL or falls exhibit a dynamic characteristic signature (*e.g.* walking, running, going upstairs but also falls) or are, on the contrary, completely static (*i.e.* lying, sitting, standing, *etc.*). Considering the whole sequence of raw data to classify a posture is thus pertinent [11]. A classical approach when working with sequences is to extract several signal feature vectors from subsequences of the signal in order to build a classifier. This approach is efficient in numerous cases but, as the window size is short, it cannot exploit long term dependencies. Moreover, some pattern recognition algorithms require the full sequence to be performed. In practice, postures vary in length: some signal are recorded over long periods of time, such as walking or lying, and others are by nature very short, such as getting up, going upstairs and falling. These differences of length could lead to two side effects when working with GRU, which could limit their capacity to learn and to generalize this learning. Firstly, sometimes, the sequence is too short to be learned properly as it does not contain enough information. Secondly, when back propagating the error through time the error, longer sequences have more influence in the learning process. The BPTT takes indeed into account every step of the sequence and it is thus performed more often for longer ones.

Therefore, we propose to resample each sensor signal of each activity sequence using a polynomial resampling setting them to equal length. We make the hypothesis that resampling the dataset could lead to better results and accelerate the learning phase (H1). Improvements could especially be observed for classes with few examples and short sequences (for instance, falls). This operation presents the disadvantage that the dataset needs to be segmented into activity sequences whereas GRU could normally classify activities on the full sequence without segmentation (*i.e.* it produces an output at each time step). Another aspect of the ADL prediction from sequences is that, considering an activity, the associated sequence can be very user specific. For example, when sleeping, every body tends to prefer different positions and move according to personal patterns. Likewise, the gait can already be used to recognize a person [17]. GRU can learn to detect these patterns and we assume that it can produce more accurate predictions when trained

only with data coming from one user (H2). Defining rules for each user is indeed time consuming and requires updating rules for each new user or each new posture category. In comparison, data-driven models are directly built without *a priori* knowledge.

IV. EXPERIMENTS

A. Datasets

1) *Postures Dataset*: The first dataset we used is the so called *Postures* dataset created by Quach [18]. The data has been acquired using a 9-axis Inertial Measurement Unit (accelerometer, gyroscopes and magnetometer, IMU) on 9 subjects executing the same sequence several times. Each user produced 5 sequences apart from user 2 who did 10. The sequences are composed of five activities or postures: walking, sitting, laying, standing and transfer and have been conceived to reflect daily routine of 24h on 24 minutes. Transfer represents the transition between two postures. Overall, there are about 358k labelled 9D vectors and 250 subsequences of vectors related to one activity. The sampling frequency is about 10 Hz. The sensor is a *Shake SK6* [19] with the following range and precision for each sensor. The range of the triple axis accelerometer is at most $\pm 6g$ with a precision of 1mg. The range of the triple axis gyroscope is of $\pm 500^\circ/s$ with a precision of 0.1 deg/second. The triple axis magnetometer has a range of ± 2 Gauss and a precision of 1 mGauss.

2) *MobiAct Dataset*: The *MobiAct* dataset [9] is an inertial dataset created to support research in ADL recognition. It includes 15 different types of labels: 4 falls and 11 ADL. The activities were recorded following a realistic scenario: a typical day of work by 67 subjects. In total, the scenario was performed around 3200 times. Data were acquired using a smartphone which the user could place anywhere. The IMU is composed of a *LSM330DLC* itself composed of a triaxis gyroscope and a triaxis accelerometer. The measurement range of the accelerometer can be selected between $\pm 2g$, $\pm 4g$, $\pm 8g$ or $\pm 16g$. The measurement range of the gyroscope can be selected between $\pm 250^\circ/s$, $\pm 500^\circ/s$ or $\pm 2000^\circ/s$. The orientation sensor combines data from the accelerometer and the magnetometer.

B. Experimental Setups

1) *Model parameters*: In our experiments, only GRU with few a few layers and neurons have been used. We defined two specific architectures based on preliminary experiences. We use the log likelihood loss and tuned our networks by using dropout and weight decay to improve the generalization. Dropout [20] randomly drops units with a probability 0.5. Weight decay adds a small penalty to the loss function for the magnitude of the weights, improving generalization [21]. Finally, during personalized training, we do not perform a validation phase as not enough data are available to constitute a validation set. Instead, we chose

a model architecture which gave reasonable results on the complete dataset and fixed the number of epochs (*i.e.* 150 epochs). Finally, each training starts with a learning rate of 0.01 which decreases by a factor 10 if the loss does not diminish during 10 epochs.

2) *Training details*: On the *Postures* Dataset, we produced personalized models for each user. The models were trained on 4 sequences (*i.e.* 9 for user 2) during 150 epochs and were tested on 1 randomly chosen sequence. Based, on preliminary experiments on the full dataset, we use personalized GRU models with two hidden layers of size 8. The process was reproduced excluding a different sequence each time (5 times for each user except for user 2, 10 times). We so perform a *k*-folds leave-one-out test of our architecture where *k* is the number of sequences associated to one user.

On *MobiAct*, we followed the same protocol as in [12] and trained our network on users 1 to 47 for training, 48 to 57 for validation and 58 to 67 for test. We let the training continue until the performance on the validation set did not increase any more during 150 epochs and then performed a test using the best parameters obtained. We use a network of 2 hidden layers of 10 units. We performed fall detection which is a binary classification (*i.e.* all classes of ADL against the 4 falls). We treat each trial as one activity, the one mentioned in the file name, and labeled the resampled sequence. For the personalized models, we split the sequences of each user randomly into two groups which was alternatively used as training and testing sets. Only a few dozens of sequences are available per user.

C. Results

1) *Personalized Posture Classification on Postures Dataset*: Makni *et al.* [13] compared two attitude device estimation algorithms and used expert rules in order to estimate the individual postures. Our first experiment consists in reproducing these experiments using only machine learning and no expert *a priori* rules. We chose also to train our models for one specific user each time. The dataset contains continuous sequences of several activities and we do not perform resampling as we cannot segment it, even on the training set. We present in the Figure 1.a the distribution of the averages of the accuracy scores obtained for all users. The boxplot shows a minimal value of 0.669, a median value of 0.732 and a maximal value of 0.874. The results are challenging for machine learning algorithms because each model is build on few inertial data sequences. This corresponds to our paradigm, where target users provide only some examples.

A comparison is presented in Table I. On average, the GRU accuracy of 0.742 is lower than the 0.807 achieved in [13]. Nevertheless, our model is not tuned for each user and no specific expert rule is applied. We used only one GRU architecture, learn on few examples and generalize well on

unknown user sequences. This is particularly encouraging when we focus on the user 2 performances which achieved the best accuracy of 0.874. This is mainly due to the fact that the user provides twice as much data as other users. Consequently, asking people to collect only 10 sequences for building a shallow GRU model is promising in practice. As a reminder, these good posture classifications would provide robustness for a eHealth actigraphy system and less parameter calibrations.

2) *Falls Detection on MobiAct*: We first remarked in [9] that falls were not as well classified as other ADL classes. Consequently, we sought to improve the performances with shallow GRU architectures, resampling data sequences and building personalized models. First, we built a generic GRU model for all users and it is thus not personalized. In Table II and Table III, we compare the GRU performances obtained with and without resampling strategy for 5 sequence sizes on the validation and test set respectively. The best validation results (*i.e.* F-measure=0.901, Accuracy=0.911) are achieved when input sequences are resampled with a length of 50. We observe a deterioration of the results when increasing the sequence length. According to us, this is due to the shallow architecture which is more adequate for compact sequences. On the test set (*cf.* Table III), the resampling 50 outperformed by 14% the GRU model without resampling, improving the F-measure value of 0.779 with 0.919. The difference is less significant with the accuracy results due to the class imbalance which affects this performance. Consequently, we conclude that resampling effectively allows better recognition of falls which validates the first part of our H1 assumption about better performances.

Secondly, we train personalized GRU models with sequences of size 50 and the same architecture [10,10], which is more realistic for building an actigraphy system, because we learn an actigraphy system only with the final user examples. Training in this configuration is very fast with around 15 seconds needed to train one model on a Quadro P6000 Nvidia GPU, which validates the second part of our H1 assumption about the learning phase acceleration. We compare, on the Figure 1.b, our GRU personalized models on the original dataset and on the resampled dataset. We observed that for 38 users out of 67, the F-measure is above 0.8, when testing personalized GRU models with resampling sequences. The F-measure is above 0.8, when testing personalized GRU models without resampling input sequences, for 32 users out of 67. Then, on average, the resampling version outperformed by 4% the GRU model without resampling, improving the F-measure value of 0.768 with 0.808 (*cf.* Table IV). Moreover, we note less variability for the resampling version and no users below 0.5. Our results are unfortunately hard to compare with those obtained in [9], as they performed their classification on the 16 classes and not only on fall detection. Nevertheless, the authors obtained with a IBk classifier the respective F-

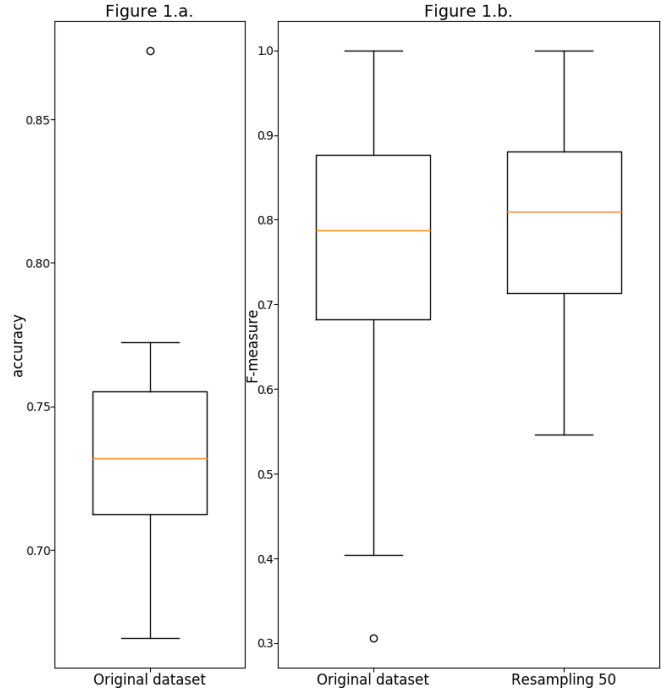


Figure 1. Average result dispersion per user for personalized models trained with our GRU architectures on Postures (a) and Mobiact V2 (b).

Table I
RESULTS FOR POSTURE CLASSIFICATION ON POSTURES WITH GRU.
(SD: STANDARD DEVIATION)

Method	F-measure	SD	Accuracy	SD
Makni et al [13]	-	-	0.807	0.024
GRU [8,8] (all users)	0.553	0.068	0.742	0.056
GRU [8,8] (user 2)	0.705	0.059	0.874	0.049

Table II
VALIDATION RESULTS FOR FALL DETECTION ON MOBIACT V2 WITH AND WITHOUT RESAMPLING WITH GRU [10,10].

Method	F-measure	SD	Accuracy	SD
without resampling	0.786	0.051	0.894	0.027
with resampling 50	0.901	0.022	0.911	0.003
with resampling 100	0.899	0.032	0.899	0.032
with resampling 200	0.889	0.033	0.889	0.033
with resampling 300	0.863	0.053	0.863	0.060
with resampling 500	0.880	0.024	0.880	0.025

Table III
TEST RESULTS FOR FALL DETECTION ON MOBIACT V2 WITH AND WITHOUT RESAMPLING WITH GRU [10,10].

Method	F-measure	SD	Accuracy	SD
without resampling	0.779	0.035	0.894	0.027
with resampling 50	0.919	0.020	0.905	0.017

Table IV
TEST RESULTS FOR PERSONALIZED FALL DETECTION ON MOBIACT V2 WITH AND WITHOUT RESAMPLING WITH GRU [10,10].

Method	F-measure	SD	Accuracy	SD
without resampling	0.768	0.161	0.883	0.135
with resampling 50	0.808	0.114	0.878	0.073

measure values 0.832, 0.760, 0.709 and 0.757 for classifying the classes "BSC", "FKL", "FOL" and "STU". By contrast, their proposal requires a lot of signal processing methods to extract high dimensional feature vectors based on human expertise. Thus, our proposed method is very promising as it uses only data with 50 samples and shallow GRU personalized models and gets the F-measure score of 0.808.

V. CONCLUSIONS AND PERSPECTIVES

We presented in this paper a new scheme based on GRU neural networks for posture and fall classification from resampled sequential data. This novel strategy relies on no predefined rules and does not need any handcrafted features. Then, we showed the potential of personalized models to improve the global performance and to accelerate the deployment of the system as GRU models require only few data sequences to train.

Our perspectives are to train personalized models on MobiAct for ADL classification. Unfortunately, this dataset does not provide enough sequences to effectively train GRU models. Consequently, we will investigate how to learn from few examples (*i.e.* one-shot learning [22]) and how to build a generic model first and then use fine-tuning strategies for the final user customization.

REFERENCES

- [1] H. Pigot, B. Lefebvre, J.-G. Meunier, B. Kerhervé, A. Mayers, and S. Giroux, "The role of intelligent habitats in upholding elders in residence," in *5th international conference on Simulations in Biomedicine*, 2003, pp. 497–506.
- [2] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, "Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey," in *ARCS*. VDE, 2010, pp. 1–10.
- [3] S. Katz, A. B. Ford, R. W. Moskowitz, B. A. Jackson, and M. W. Jaffe, "Studies of illness in the aged: the index of adl: a standardized measure of biological and psychosocial function," *JAMA*, vol. 185, no. 12, pp. 914–919, 1963.
- [4] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [5] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [6] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *International Conference on Pervasive Computing*. Springer, 2004, pp. 1–17.
- [7] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," *Pattern recognition letters*, vol. 29, no. 16, pp. 2213–2220, 2008.
- [8] R. Jafari, W. Li, R. Bajcsy, S. Glaser, and S. Sastry, "Physical activity monitoring for assisted living at home," in *4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007)*. Springer, 2007, pp. 213–219.
- [9] C. Chatzaki, M. Pediaditis, G. Vavoulas, and M. Tsiknakis, "Human daily activity and fall recognition using a smartphone's acceleration sensor," in *ICT4AWE*. Springer, 2016, pp. 100–118.
- [10] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, and J. Zhang, "Convolutional neural networks for human activity recognition using mobile sensors," in *IEEE MobiCASE*, 2014, pp. 197–205.
- [11] G. Lefebvre, S. Berlemont, F. Mamalet, and C. Garcia, "BLSTM-RNN based 3D gesture classification," in *ICANN*. Springer, 2013, pp. 381–388.
- [12] R. DiPietro, C. Rupprecht, N. Navab, and G. D. Hager, "Analyzing and exploiting NARX recurrent neural networks for long-term dependencies," *Workshop track - ICLR*, 2017.
- [13] A. Makni and G. Lefebvre, "Attitude estimation for posture detection in ehealth services," in *IEEE CBMS*, 2018, pp. 310–315.
- [14] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [16] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.
- [17] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette analysis-based gait recognition for human identification," *IEEE PAMI*, vol. 25, no. 12, pp. 1505–1518, 2003.
- [18] K. A. Quach, "Extraction de caractéristiques de l'activité ambulatoire du patient par fusion d'informations de centrales inertielles," Ph.D. dissertation, UCBL1, 2012.
- [19] J. Williamson, R. Murray-Smith, and S. Hughes, "Shoogle: excitatory multimodal interaction on mobile devices," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007, pp. 121–124.
- [20] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *JMLR*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [21] A. Krogh and J. A. Hertz, "A simple weight decay can improve generalization," in *NIPS*, 1992, pp. 950–957.
- [22] D. Wu, F. Zhu, and L. Shao, "One shot learning gesture recognition from RGBD images," in *Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2012, pp. 7–12.