A Hierarchical Approach to Recognize Purposeful Movements Using Inertial Sensors: Preliminary Experiments and Results

Carme Zambrana
Eurecat Technology Center,
eHealth Unit
Barcelona, Spain
carme.zambrana@eurecat.org

Sebastian Idelsohn-Zielonka
Eurecat Technology Center,
eHealth Unit
Barcelona, Spain
sebastian.idelsohn@eurecat.org

Mireia Claramunt-Molet
Eurecat Technology Center,
eHealth Unit
Barcelona, Spain
mireia.claramunt@eurecat.org

Maria Almenara-Masbernat
Institut Guttmann,
Neurorehabilitation Institute
Badalona, Spain
malmenara@guttmann.com

Eloy Opisso
Institut Guttmann,
Neurorehabilitation Institute
Badalona, Spain
eopisso@guttmann.com

Josep Maria Tormos
Institut Guttmann,
Neurorehabilitation Institute
Badalona, Spain
jmtormos@guttmann.com

Felip Miralles
Eurecat Technology Center,
eHealth Unit
Barcelona, Spain
felip.miralles@eurecat.org

Eloisa Vargiu
Eurecat Technology Center,
eHealth Unit
Barcelona, Spain
eloisa.vargiu@eurecat.org

ABSTRACT
One of the most relevant post-stroke conditions is the hemiparesis, which causes muscle weakness and/or the inability to move one side of the body. Physical and occupational therapy plays an important role in the rehabilitation of patients suffering this condition. On the other hand, daily life use of the impaired arm is crucial for improving and also assessing the evolution of the patient. Currently, this assessment is done through self-questionnaires and interviews, which are subjective and depend on the memory of the patient. In this paper, a hierarchical automatic approach aimed at recognizing purposeful arm movements during patients' daily life activities is presented. This approach relies on two-levels: the former is aimed at distinguishing between arm movement and non-movement; whereas the latter is devoted to recognize between purposeful and non-purposeful movements. In particular, in the first version of the system, we consider arms swing while walking as non-purposeful movement. Experiments have been performed in the lab with 9 healthy volunteers wearing a wristband on each wrist. Six activities have been performed: eating, pouring water, drinking, brushing their teeth, folding a towel, and walking. The proposed approach achieves promising performances, recognizing purposeful movement with an accuracy of 0.91 and an F1-score of 0.87.

PervasiveHealth '17, May 23–26, 2017, Barcelona, Spain

ACM ISBN 978-1-4503-6363-1/17/05…$15.00
https://doi.org/10.1145/3154862.3154932

ACM Classification Keywords
Applied computing: Life and medical sciences: Health care information systems

Author Keywords
rehabilitation; remote monitoring; internet of things; inertial sensors.

INTRODUCTION
According to the World Health Organization[1], 15 million people suffer stroke worldwide each year. Of these, 5 million die and another 5 million are permanently disabled. As a consequence, stroke produces immense health and economic burdens globally [15].

Many of stroke survivors have paralysis and/or balance problems. Statistics show that 40% of all stroke survivors suffer serious falls within a year after their stroke [2]. Hemiparesis or one-sided ("hemi") weakness ("paresis") affects about 8 out of 10 stroke survivors, causing weakness or the inability to move one side of the body [10]. Hemiparesis can affect arms, hands, legs, and facial muscles. Thus, people that have

[1] http://www.who.int/en/

hemiparesis may have trouble performing everyday activities such as eating, dressing, or toileting.

Physical and occupational therapy plays an important role in the rehabilitation of patients suffering hemiparesis [19, 20]. Moreover, rehabilitation treatments, exercises at home, and assistive devices can help with mobility and recovery [5, 17]. In the literature, several solutions have been proposed to perform rehabilitation at home. Solutions range from home rehabilitation teams that helps in doing activities chosen on the basis of the patient's personal interests [9] to virtual environments and virtual reality specifically thought for this task [8].

Apart from rehabilitation in hospitals and specialized centers, physiotherapists agree that the use of the impaired limb during their normal life –i.e., performing daily life activities (ADLs)– radically improve patients' recovery. Moreover, rehabilitation can be performed at home. Assessment, in this case, is left to the patients that through specific questionnaires report on the performed activities [16]. Unfortunately, this kind of solution is quite subjective and depends totally on the specific patients, her/his feeling, and also her/his cognitive status. Thus, automatic solutions able to identify performed ADLs and the use of the impaired limb are required.

This paper presents a first solution in that direction. Focusing on upper-limp impairment, it proposes a hierarchical approach aimed at recognizing purposeful movements performed during patients' ADLs with both arms. It is composed of two-levels that work sequentially. First, the upper-level module distinguishes between arm movements and non-movements. Then, data filtered by that module are analyzed by the lower-level module that recognizes between purposeful and non-purposeful movements. It is worth noting that in this first version, only arms swing while walking has been considered as non-purposeful movement. With the aim to move this solution at stroke patients' home, first experiments have been performed in the lab with 9 healthy volunteers wearing a wristband on each wrist. Besides walking (non-purposeful movement), five activities in which the arm has to be purposefully moved have been performed: eating; pouring water; drinking; brushing their teeth; and folding a towel.

The rest of the paper is organized as follows: first, the background is presented to put in the context the proposed approach and its novelty. Subsequently, the approach is described focusing on the technologies adopted in each of its levels. Then, preliminary experiments and results are presented and discussed. Finally, the paper ends with conclusions and directions for future work.

BACKGROUND

As stated in [6], purposeful versus non-purposeful movement has generated a great deal of debate among rehabilitation researchers. On the one hand, researchers claim the relevance to quantify only purposeful movements, which are part of a focused upper-limb related task, thus considering swing as a non-purposeful movement. On the other hand, there are plausible reasons to consider all kinds of motions, including swing, as purposeful. Thus, there is a strong clinical rationale for including all kinds of movements, both purposeful and non-purposeful, in ADLs recordings.

Several research studies address this important issue of recognizing purposeful and non-purposeful movements. Being interested on arm movements, let us recall here some relevant work. In [14], authors consider as non-purposeful movements arm movements caused by whole-body movements, and as purposeful movements all of those that are independent of whole-body movements. Uswatte et al. [22] define as purposeful movement (referred as "functional activity") every movement or action that helps to accomplish a task (e.g., grasping a can and wiping a table top) or, more generally, that has some specific functionality (e.g., touching a face and moving the arm from one position to another). On the other hand, they define as non-purposeful movements (referred as "non-functional activity") movements or actions that are not correlated with any functionality (e.g., tic and tremor) or that are largely secondary to movement of other parts of the body (e.g., arm swing while walking and passive movement). On the contrary, in [1], authors claim that deciding what constitutes non-functional movement (e.g., a tick or jerk) during quiescent periods is subjective; several authors try to define how purposeful and non-purposeful movements may be defined.

Similarly to our approach, in [21] authors consider functional arm movements (e.g., getting dressed) as purposeful movements and non-functional (e.g. arm swing while ambulating) as non-purposeful. Seemly, McLeod et at. [13] consider walking as non-functional activity and they refer as non-purposeful movement the swing during that.

From a technological perspective, several solutions that rely on wearable devices and in particular on wristbands have been proposed for recognizing activities as well as differentiating between purposeful and non-purposeful ones. In [11], authors studied how to quantify functionally arm use in stroke survivors monitoring patients for 48 hours. They used 5 worn IMUs (with 3-axis accelerometer, gyroscope, magnetometer and a barometric pressure sensor); placed 2 on the wrists, 2 on the shanks and 1 on the waist. Data coming from the shank-worn sensors where used to identify the walking activity. Then, the activity counts where calculated using the wrist-worn sensors. On the contrary, with the aim to be less invasive, our approach only uses wrist-worn sensors for distinguish the walking activity from others, relying on machine learning techniques for classifying them. In [13], custom IMU devices were used with 3 orthogonally positioned linear accelerometers, 3 orthogonally positioned gyroscopes, and a 3-dimensional magnetometer. Only the 3 linear accelerometers and 3 angular rate sensor data channels were retained. Magnetometer data were discarded because of indoors unreliability and the fact that cardinal orientation of the subjects during performance was fixed for all trials. Sensors were attached to the dominant wrist of controls and the prosthetic wrist of amputees. Data were collected in a hospital training apartment. The following activities were monitored: clothes folding; bot stacking; carrying; and walking. Seemly to us they consider walking activity as nonfunctional activity. But their approach consists of a flat machine learning model whereas

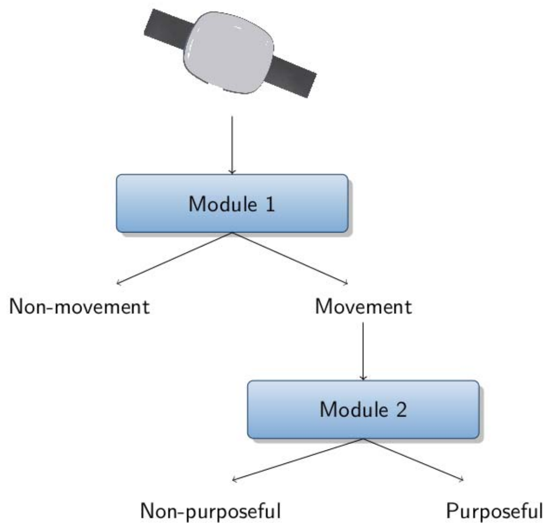

Figure 1. The proposed hierarchical approach.

we propose a hierarchical approach for its scalability to other non-purposeful activities.

METHOD

With the aim of recognizing purposeful and non-purposeful movements during patient's ADLs, we propose an automatic hierarchical approach. Generally speaking, a hierarchical approach combines different modules linked together by a hierarchical relationship. In so doing, a complex problem can be decomposed in easier sub-problems according to a divide-et-impera approach.

The proposed hierarchical approach is shown in Figure 1: the upper level (Module 1) is aimed at distinguishing between arm movements and non-movements; whereas the lower one (Module 2) is devoted to recognize between purposeful and non-purposeful movements of the arms.

Module 1: Movement versus Non-Movement

Taking as input the raw data from a wristband, the aim of Module 1 is to distinguish between samples that correspond to a movement from those that correspond to non-movements. First, a pre-processing phase is applied to the overall signal. Then, the best threshold is searched and the results refined using the Tartaglia's triangle filter.

The pre-processing phase consists of combining the 3 signals of the tri-axial accelerometer by using the Signal Magnitude Vector (SMV) [12]:

$$SMV = \sqrt{acc_x^2 + acc_y^2 + acc_z^2}$$

Starting from the SMV of each sample, the best threshold θ is calculated. It is then used to classify each sample as movement or non-movement, depending on whether the value is higher (*Movement*) or lower (*Non-Movement*) than θ. Results are then refined. In fact, each sample is reclassified by taking into account the class that has been assigned to its 4 nearest neighbors (the 2 previous and the 2 posterior ones). Neighbors are weighted using the Tartaglia's triangle: $[1 \;\; 3 \;\; 3 \;\; 1]$ $(= [f_{-2} \;\; f_{-1} \;\; f_1 \;\; f_2])$. For each sample its value v_i is calculated according to the formula:

$$v_i = \sum_{q=-2(\neq 0)}^{2} l_{i+q} \cdot f_q$$

where l_i is the label corresponding to the i-th sample: 0 for *Non-Movement* and 1 for *Movement*. Moreover, in order to be able to calculate the value of the 2 first and the 2 last samples, 4 *Non-Movement* labels have been included: l_{-2}, l_{-1}, l_{n+1}, $l_{n+2} = 0$. Finally, if the value is bigger than half of the maximum value (v_{max}), the sample is reclassified as *Movement*, otherwise as *Non-Movement*:

$$v_{max} = \sum_{q=-2(\neq 0)}^{2} 1 \cdot f_q = 8$$

It is worth noting that the Tartaglia's triangle filter is applied sequentially, which means that each sample is classified according to the previous 2 neighbors refined by using the Tartaglia's triangle filter and the following 2 neighbors.

The output of Module 1 is the set of samples recognized as *Movement* according to the Tartaglia's triangle filter.

Module 2: Purposeful Movement versus Non-Purposeful Movement

The aim of Module 2 is to recognize between purposeful and non-purposeful movements. Let us stress that in the first version presented in this paper, we consider as non-purposeful movements arms swing while walking.

Taking as input the SMV of the samples recognized as *Movement* by Module 1, Module 2 uses a supervised binary classifier to distinguish between purposeful and non-purposeful movements. Let us recall here that a supervised binary classifier is aimed at classifying samples in 2 given classes (*binary*). That classifier is built by relying on a training set that contains samples whose class is known (*supervised*).

The output of Module 2 is the class given by the classifier to each sample.

PRELIMINARY EXPERIMENTS AND RESULTS

Experiments have been performed in the lab with 9 healthy volunteers (mean age 31.22 ± 4.59 years, 5 women and 4 men, 1 left-handed) wearing a wristband on each wrist. Six activities have been performed: eating, pouring water, drinking, brushing their teeth, folding a towel, and walking.

Devices

Experiments have been performed with 2 IMUs, including a tri-axial accelerometer, a tri-axial gyroscope and two tri-axial

Figure 2. The adopted wristband and the orientation of the tri-axial accelerometer.

magnetometers, located on both wrists. The frequency acquisition of these devices is 20 Hz and the data are transmitted via Bluetooth. In these preliminary experiments only the data from accelerometer have been used. Figure 2 sketches the orientation of the tri-axial accelerometer.

Activities

Together with physiotherapists from the Institut Guttmann, the protocol to be used during the experiments has been defined. In fact, let us note that even if in this paper we are reporting results with healthy volunteers, the main aim of the underlying overall project is to give assistance to stroke survivors. Thus, the protocol used in the experiments presented in this paper will be adopted to test the approach with people with an impaired arm.

The adopted protocol is the following:

1. *eating*. A knife, a fork and a piece of play-dough, simulating a hamburger, were on the table. Volunteers were instructed to simulate eating the hamburger;

2. *pouring water*. An empty cup and a bottle with water were on the table. Volunteers were instructed to pour water into the cup;

3. *drinking*. A cup with water was on the table. Volunteers were instructed to drink the water;

4. *brushing their teeth*. A toothbrush was on the table. Volunteers were instructed to simulate brushing their teeth (without toothpaste);

5. *folding a towel*. An unfolded towel was on the table. Volunteers were instructed to fold the towel, leave it on the table for a few seconds, and unfold it;

6. *walking*. Volunteers were instructed to walk through a 12-meter-long corridor.

Before performing each activity, the researcher explained to the volunteers the activity to be performed, but not how to do it. The volunteers were instructed to perform activities from 1 to 5 three times, resting 10 seconds between each repetition. For the sake of creating a balanced dataset, activity 6 was recorded during 4 minutes.

Figure 3. The markers put on the wristbands to be tracked through the Trio-OptiTrack.

Module 1: Movement versus Non-Movement

Actions from 1 to 5 were recorded by a Trio-OptiTrack[2]. It is composed of 3 infrared cameras capable of tracking at 120 fps the reflective markers placed on the device (see Figure 3). Data kept by the Trio-OptiTrack have been used to create the ground truth for the Module 1 (i.e., to distinguish between movements and non-movements). These data consist of the absolute position of the markers. To define the reference value of non-movement each volunteer was asked to remain in the resting position for 30 seconds (with both forearms on the table, as shown in Figure 3). The idea is to use for each sample the relative position between the current position and the previous one, thus labeling it as *Movement* when the position differs more than a given distance d. This distance has been computed using data belonging to the resting activity. For each volunteer, the maximum distance between two consecutive samples of resting has been used to calculate d. The walking activity has not been recorded in this way, because it exceeds the maximum capture volume of the cameras. Thus, it has been manually labeled by using video recording.

As stated above, the raw data from the accelerometer have been pre-processed by using the SMV. The best threshold (θ) has been obtained performing a grid search over the range from 0 to 1 by 0.005 using the data coming from actions from 1 to 5. The adopted metric used to report the performance by comparing the obtained labels (after applying the threshold and the *Tartaglia*'s triangle filter) and the ground truth is the accuracy:

$$Acc = \frac{TP+TN}{TP+TN+FN+FP}$$

where TP is the number of True Positives, i.e., samples corresponding to movements that have been recognized as *Movement*; TN is the number of True Negatives, i.e., samples corresponding to non-movements that have been recognized as *Non-Movement*; FN is the number of False Negatives, i.e.,

[2]http://optitrack.com/products/v120-trio/

θ	Acc
0.365	0.89204
0.355	0.89201
0.360	0.89193
0.385	0.89170
0.350	0.89141

Table 1. Module 1: top-5 results, for each threshold θ the measured accuracy is given.

samples corresponding to movements that have been recognized as *Non-Movement*; and FP is the number of False Positives, i.e., samples corresponding to non-movements that have been recognized as *Movement*.

Table 1 reports the top-5 thresholds. As shown, the best threshold is $\theta = 0.365$, which gives an accuracy $Acc = 0.89204$ over the data coming from actions 1 to 5. With this threshold the overall accuracy of Module 1 is $Acc = 0.94270$.

Module 2: Purposeful vs Non-Purposeful

Samples recognized as *Movement* by Module 1 are the input of Module 2. In particular, the corresponding dataset contains, for each sample, the SMV and the time stamp of when it was recorded.

First, two sets of features have been considered: time-domain and frequency-domain, respectively [12]. The time-domain features are: mean; standard deviation; minimum value; maximum value; and peak-to-peak amplitude. The frequency-domain features are: total power; the dominant frequency and its power; the second-dominant frequency and its power; the dominant frequency and its power for the band $[0.3\text{-}10]Hz$; and the ratio between dominant frequency power and total power.

These features have been computed using fixed-length, non-overlapping windows. According to Mannini et al. [12], we considered windows of 2.0 and 4.0 seconds[3].

Three state-of-the-art binary supervised classifiers have been experimented: k-Nearest Neighbor (KNN), Random Forest (RF), and Support Vector Machine (SVM). KNN [4] is a simple algorithm that uses the k closest training samples in the feature space to predict the class of a new given sample. Distance between samples is calculated using the Euclidean distance and the predicted class is the most common. A weight can be added to the contributions of the neighbors, so that the nearest neighbors contribute more to the average than to the actual distance. The number of neighbors (k) should be a positive integer. In binary classification, when the neighbors contribute to the same weight, the k should be an odd number, avoiding possible ties. RF [7] is an averaging ensemble learning method. This method builds several estimators independently and then averages their predictions. Each of these estimators is a Decision Tree, which learns simple decision rules inferred from the data features [18]. The size of the random subset of features considered when splitting a node can be delimited by fixing its maximum number of features in the subset. The SVM [3] is the most complex algorithm of the three considered for experiments. In the binary case, the machine conceptually classifies the samples by using a hyperplane. When the training set is not perfectly separable, the regularization parameter (C) determines the weight of the samples misclassified. The shape of the hyperplane is defined by a kernel, which could be, among others, linear or Gaussian, also known as Radial Basis Function (RBF). Using RBF kernel the parameter γ could be used to define the influence of the training samples.

[3]Windows of 12.8 were not taken into consideration because of the length of the performed activities.

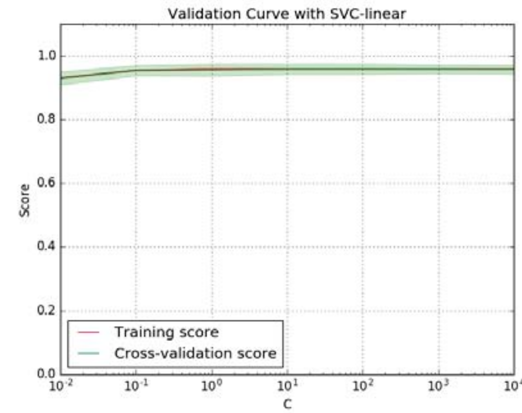

Figure 4. Module 2: Validation curve of the best classifier.

The dataset has been split in two sub-datasets: one for training and one for testing (70%-30%). Each supervised classifier has been built from the training set and a grid search 10-fold cross validation has been used in order to find the best set of parameters. Cross-validation is a well-known method which consists of splitting the training set in k-folds, then each of this partitions is used as test while the model is trained using the remaining $k-1$, the final score is obtained by averaging these k scores.

Experiments have been performed to find the best configuration of the parameters; i.e. the set of features, the length of the windows, as well as the best classifier. Table 2 shows the scores in accuracy obtained by the cross-validation method over the training sets.

The best fours classifiers, one for each dataset, are tested over the test sets, Table 3 shows the results including the F1-score obtained over the minor class and the total averaged F1-score (minor class/total):

$$F1 = 2 \cdot \frac{TP}{2TP + FN + FP}$$

As shown in Table 3, the best performance has been obtained considering time-domain features computed with 2.0-seconds windows and for the Linear-SVM with $C = 10$; the corresponding F1-score being $F1 = 0.95699$ The validation curve of this classifier is presented in Figure 4. As shown, the classifier does not over-fit over the training set.

	Time domain		Frequency domain	
	2	4	2	4
SVM-Linear (C)	**(10) 0.95857 ± 0.01673**	(1.0) 0.97987 ± 0.0194	(1.0) 0.86791 ± 0.02367	(10) 0.90021 ± 0.02798
SVM-RBF $\gamma = \frac{1}{numfeatures}$ (C)	(1.0) 0.95595 ± 0.016193	**(10) 0.98130 ± 0.02127**	(10) 0.90755 ± 0.02447	(10) 0.93366 ± 0.02297
SVM-RBF C = 1.0 (γ)	(0.1) 0.95418 ± 0.01833	(0.1) 0.97840± 0.01835	(1.0) 0.90842 ± 0.02074	(1.0) 0.940756 ± 0.02176
RF max features = $\sqrt{numfeatures}$ (num estimators)	(32) 0.95155± 0.01256	(26) 0.97845 ± 0.02139	(44) 0.91371 ± 0.02056	(29) 0.94664 ± 0.03012
RF num estimators = 10 (max features)	(4) 0.94979 ± 0.01961	(4) 0.96832 ± 0.01991	**(3) 0.91899 ± 0.01733**	**(5) 0.94807 ± 0.02032**
KNN-uniform (num neighbors)	(7) 0.95242 ± 0.01632	(7) 0.97987 ± 0.02043	(11) 0.90663 ± 0.02759	(5) 0.93492 ± 0.02851
KNN-distance (num neighbors)	(7) 0.95244 ± 0.01528	(7) 0.97987 ± 0.02043	(13) 0.90488 ± 0.0269	(5) 0.93639 ± 0.02990

Table 2. Module 2: the cross-validation results obtained for each set of features and each length of the window. Results are given in terms of the mean of the accuracy and its standard deviation.

	Window length	
	2.0	*4.0*
Time	**SVM(Linear, 10)** **0.94988 / 0.95699**	SVM (RBF, 10, 0.2) 0.94468/0.95645
Frequency	RF (10, 3) 0.89604/0.91357	RF (10, 5) 0.92241/0.93958

Table 3. Module 2: the best classification results obtained for each set of feature and each length of window. Results are given in terms of the F1-score obtained over the minor class / the total averages F1-score.

Overall Results

Finally, to calculate the overall results, we tested the hierarchical approach with the best configuration of Module 1 ($\theta = 0.365$) and Module 2 (Linear-SVM with time-domain features and a window of 2.0 seconds). The approach shows high accuracy and F1-score in recognizing purposeful movements: $Acc = 0.90701$ and $F1 = 0.86500$.

CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a two-level hierarchical approach to recognize purposeful movements of upper-limbs using inertial sensors. The upper-level of the hierarchy (Module 1) is aimed at recognizing between movement and non-movement; whereas the lower-level (Module 2) is aimed at distinguishing between purposeful and non-purposeful movements. Preliminary experiments have been performed with 9 healthy volunteers and results are very promising. In fact, results of Module 1 reached an accuracy of 0.94270. To calculate results of Module 2, four set of experiments have been performed by combining time domain features and frequency domain features with 2 fixed-length windows. An F1-score of 0.95699 has been obtained by using the time-domain features with a window of 2.0 seconds with an Linear-SVM. Results of the overall hierarchical approach show an accuracy of 0.90701 and an F1-score of 0.86500. These results are very encouraging. It is worth noting that instead of a hierarchical approach, a flat one able to recognize purposeful movements could be investigated and developed. Nevertheless, let us stress that hierarchically dividing the problem reduces the computational cost of the corresponding system. Being the final goal to embed –totally or partially– the hierarchical approach presented in this paper within a wearable device, computational cost is indeed a relevant factor. In particular, we are currently working to embed Module 1 in the adopted wristband.

As for the future work, we are currently setting up the experiments to validate the clinical relevance of the proposed hierarchical approach. This validation includes a pilot study with post-stroke patients performing the ADLs presented in this paper. The pilot will be carried out in a rehabilitation center specialized in neurological disabilities. As a further next step, we are researching how to define specific modules aimed at recognizing other non-purposeful movements.

ACKNOWLEDGMENTS

The study has been partially funded by ACCIÓ (Pla d'Actuació de Centres Tecnològics 2016) under the project ADALT.

REFERENCES

1. Ryan R Bailey and Catherine E Lang. 2014. Upper extremity activity in adults: referent values using accelerometry. *Journal of rehabilitation research and development* 50, 9 (2014), 1213.
2. Fiona M Collen, Derick T Wade, and Carole M Bradshaw. 1990. Mobility after stroke: reliability of measures of impairment and disability. *International disability studies* 12, 1 (1990), 6–9.
3. Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine learning* 20, 3 (1995), 273–297.
4. Thomas Cover and Peter Hart. 1967. Nearest neighbor pattern classification. *IEEE transactions on information theory* 13, 1 (1967), 21–27.
5. Ulla-Britt Flansbjer, Anna Maria Holmbäck, David Downham, Carolynn Patten, and Jan Lexell. 2005. Reliability of gait performance tests in men and women with hemiparesis after stroke. *Journal of rehabilitation medicine* 37, 2 (2005), 75–82.
6. Kathryn S Hayward, Janice J Eng, Lara A Boyd, Bimal Lakhani, Julie Bernhardt, and Catherine E Lang. 2016. Exploring the Role of Accelerometers in the Measurement of Real World Upper-Limb Use After Stroke. *Brain Impairment* 17, 01 (2016), 16–33.
7. Tin Kam Ho. 1995. Random decision forests. In *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, Vol. 1. IEEE, 278–282.
8. Maureen K Holden. 2005. Virtual environments for motor rehabilitation: review. *Cyberpsychology & behavior* 8, 3 (2005), 187–211.

9. L Widen Holmqvist, L Von Koch, V Kostulas, M Holm, G Widsell, H Tegler, K Johansson, J Almazan, and J de Pedro-Cuesta. 1998. A randomized controlled trial of rehabilitation at home after stroke in southwest Stockholm. *Stroke* 29, 3 (1998), 591–597.

10. Rashmi Kothari, Laura Sauerbeck, Edward Jauch, Joseph Broderick, Thomas Brott, Jane Khoury, and Tiepu Liu. 1997. Patients? awareness of stroke signs, symptoms, and risk factors. *Stroke* 28, 10 (1997), 1871–1875.

11. Kaspar Leuenberger, Roman Gonzenbach, Susanne Wachter, Andreas Luft, and Roger Gassert. 2017. A method to qualitatively assess arm use in stroke survivors in the home environment. *Medical & biological engineering & computing* 55, 1 (2017), 141–150.

12. Andrea Mannini, Stephen S Intille, Mary Rosenberger, Angelo M Sabatini, and William Haskell. 2013. Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise* 45, 11 (2013), 2193.

13. Adam McLeod, Elaine M Bochniewicz, Peter S Lum, Rahsaan J Holley, Geoff Emmer, and Alexander W Dromerick. 2016. Using Wearable Sensors and Machine Learning Models to Separate Functional Upper Extremity Use From Walking-Associated Arm Movements. *Archives of physical medicine and rehabilitation* 97, 2 (2016), 224–231.

14. Marian E Michielsen, Ruud W Selles, Henk J Stam, Gerard M Ribbers, and Johannes B Bussmann. 2012. Quantifying nonuse in chronic stroke patients: a study into paretic, nonparetic, and bimanual upper-limb use in daily life. *Archives of physical medicine and rehabilitation* 93, 11 (2012), 1975–1981.

15. Dariush Mozaffarian, Emelia J Benjamin, Alan S Go, Donna K Arnett, Michael J Blaha, Mary Cushman, Sarah De Ferranti, Jean Pierre Després, Heather J Fullerton, Virginia J Howard, and others. 2015. Executive summary. *Circulation* 131, 4 (2015), 434–441.

16. FM Nouri and NB Lincoln. 1987. An extended activities of daily living scale for stroke patients. *Clinical rehabilitation* 1, 4 (1987), 301–305.

17. Michele Pirovano, Renato Mainetti, Pier Luca Lanzi, and Nunzio Alberto Borghese. 2014. Game Engines and Exergames to Guide Rehabilitation at Home. In *Replace, Repair, Restore, Relieve–Bridging Clinical and Engineering Solutions in Neurorehabilitation*. Springer, 129–134.

18. J. Ross Quinlan. 1986. Induction of decision trees. *Machine learning* 1, 1 (1986), 81–106.

19. ME Smith, WM Garraway, DL Smith, and AJ Akhtar. 1982. Therapy impact on functional outcome in a controlled trial of stroke rehabilitation. *Archives of physical medicine and rehabilitation* 63, 1 (1982), 21–24.

20. A Sunderland, DJ Tinson, EL Bradley, D Fletcher, R Langton Hewer, and DT Wade. 1992. Enhanced physical therapy improves recovery of arm function after stroke. A randomised controlled trial. *Journal of Neurology, Neurosurgery & Psychiatry* 55, 7 (1992), 530–535.

21. Gitendra Uswatte, Carol Giuliani, Carolee Winstein, Angelique Zeringue, Laura Hobbs, and Steven L Wolf. 2006. Validity of accelerometry for monitoring real-world arm activity in patients with subacute stroke: evidence from the extremity constraint-induced therapy evaluation trial. *Archives of physical medicine and rehabilitation* 87, 10 (2006), 1340–1345.

22. Gitendra Uswatte, Wolfgang HR Miltner, Benjamin Foo, Maneesh Varma, Scott Moran, and Edward Taub. 2000. Objective measurement of functional upper-extremity movement using accelerometer recordings transformed with a threshold filter. *Stroke* 31, 3 (2000), 662–667.