Infusing Domain Knowledge to Improve the Detection of Alzheimer's Disease from Everyday Motion Behaviour

Chao Bian[1], Shehroz S. Khan[⋆2], and Alex Mihailidis[12]

[1] chao.bian@mail.utoronto.ca
Institute of Biomaterials and Biomedical Engineering,
[2] {shehroz.khan, alex.mihailidis}@utoronto.ca
Department of Occupational Science and Occupational Therapy,
University of Toronto, Canada.

Abstract. Alzheimer's disease can severely impair the independent lifestyle of a person. Dem@Care is an European research project that conducted a study for timely diagnosis of Alzheimers disease by collecting everyday motion data from couples (or dyads), with one of the person in the couple having AD. Their results suggest that AD can be detected using everyday motion data from accelerometers. They did evaluation based on leave-one-person-out cross-validation. However, this evaluation can introduce bias in the classification results because one of the person from the dyad is present in the training set while the other is being tested. In this paper, we revisit the Dem@Care study and propose a new evaluation method that performs leave one-dyad-out cross-validation to remove the dataset selection bias. We then introduce new domain specific features based on dynamic and static intervals of motions that significantly improves the classification results. We further show increase in performance by combining the proposed features with new time, frequency domain and baseline features used in the Dem@Care study.

Keywords: machine learning, cross-validation, accelerometer, Alzheimer's disease, feature extraction

1 Introduction

As the population of elderly people increase, more cases of Alzheimer's disease (AD) and dementia are reported. According to Alzheimer's Association, one in 10 Americans age 65 and older has AD in 2017 [1]. AD is listed as the sixth-leading cause of death in the United States [1]. AD destroys brain cells that impairs the thinking ability and deteriorates memory, which in turn impairs everyday activities of daily living (ADL) by significantly changing the temporal structure of these activities [2]. An early detection and prediction of such behaviours may allow the application of required interventions to deal with this

⋆ Corresponding Author

cognitive degeneration problem. However, typical clinical assessment approaches have an episodic nature [3] and tend to introduce a high level of subjectivity [4]. Clinicians also lack a comprehensive view of the person's ADL, which can only be provided by continuous monitoring [4].

With the advancements in sensor technology and cost effective sensing devices, it is feasible to perform continuous and longitudinal monitoring of elderly people in smart homes [3, 2]. Dem@Care is an European research project for timely diagnosis, assessment, maintenance and promotion of self-independence of people with dementia [5]. One of their studies is related to detecting the effects of AD from everyday motion behavior [2]. This study used machine learning methods on motion data collected from community-dwelling individuals with and without AD. Their results suggest that changes of everyday behavior in people with AD are detectable in motion data collected using accelerometer. They collected motion data from several dyads (couples) with one partner diagnosed with AD and one partner being a healthy control. They extracted frequency domain and principal component analysis (PCA) features from this data, tested several machine learning algorithms using leave-one-subject-out cross-validation (LOSOCV) and reported high classification accuracy.. The LOSOCV methodology adopted in this study suffers from a drawback that one participant of the dyad is always present in the training set while the other participant of that dyad is being tested. Since both the participants in a dyad lived together, they may bear resemblance in their lifestyle and ADL. During the LOSOCV, an inadvertent bias may be introduced while training the model leading to high detection rates. In such a case, the classifier could be mapping a 'similar' participant present in the training set to a participant being tested, rather than identifying the patterns related to AD. To circumvent the problems associated with the above cross-validation method, we propose to use leave-one-dyad-out cross-validation (LODOCV) that keeps the dyad for testing separate from the training test . Then, we utilize domain knowledge about the motor behaviour of people with AD and infuse that knowledge to the classifiers with corrected evaluation method. Our results show improved detection rate by combining motion based domain features with new evaluation methodology.

The rest of the paper is structured as follows. Section 2 presents literature review on using sensor technology and machine learning techniques to identify AD or dementia. In Section 3, we describe the Dem@Care dataset, data preprocessing, and feature extraction methods. In Section 4, we present the new dynamic and static interval based features along with new time and frequency features to handle this problem. Section 5 presents the classification results under both LOSOCV and LODOCV framework and different feature regimes. Section 6 concludes the paper with pointers to future research directions.

2 Literature Review

Researchers have been using various sensors and machine learning techniques to detect dementia in older adults. Accelerometer is the most commonly used body-worn sensor for dementia detection [6, 7]. Researchers have also adopted heterogeneous sensors to create smart ambient environment, or smart homes, for assessing cognitive health and detecting dementia related symptoms. For instance, passive infrared motion sensors were installed in home environments to detect motion patterns of older adults at home [8]. Sensors were attached on objects or items at home such as door, telephone, broom, tea cup, kettle, etc., to monitor activities that involved interaction with these objects [9, 10]. Wristband that contains sensors to monitor motion intensity and vital signs were also used in several research [8, 11, 3]. Vision sensors that are either in the form of camera or kinect are also used for this task [4, 12, 13].

The aforementioned sensors used in dementia detection generated heterogeneous data, which are used to extract features of monitored targets. Examples of those features include motion, physical activity, use of appliances, etc. Normally, the extracted features from sensor data are either statistical, temporal, frequency based, or domain specific. For instance, features extracted from accelerometers includes zero crossing rate [14, 6], area under magnitude of the Fast Fourier Transform (FFT) [15], peak acceleration and energy [16], etc. Kinect-based vision sensor provided features like relative and absolute joint angles and distance to the hip center [17]. Smart home sensors usually generate event-based or triggering-based features such as binary, change-point and last-fired features [11]. These features were then fed into various algorithms for dementia detection. There were mainly two type of algorithms, namely, statistical [18] and machine learning [3, 19, 9]. Commonly used statistical methods were analysis of variance and correlation analysis. These methods were used to study the correlations between sensor data and dementia-related parameters such as apathy [20], agitation [14] and gait [21]. Various machine learning algorithms have been used to detect dementia-related symptoms or abnormal activities, for example, Support Vector Machine algorithm has been used for agitation detection [15, 16], mobility assessment [22] and activities recognition [23].

Most of the studies were conducted in an instrumented lab environment [22, 17], however, a free-living environment may be more naturalistic to test the real effect of proposed methods [23, 2]. The study by Kirste et al. [2] involved two partners (a dyad) living in the same residence/home. In each dyad, one partner was diagnosed with AD and the other one was a healthy control. Moreover, the number of study participants in different studies varied greatly from less than 10 to more than 100 [6, 14]. Some of the studies only recruited healthy adults and simulated behaviors of older adults based on defined protocols [3, 16]. These proposed sensor systems and algorithms reported various detection accuracy for dementia. Alam et al. [23] and Dawadi et al. [24] both investigated automated assessment of activity performed by older adults with dementia to detect function decline due to reduced task quality. Alam [23] claimed sequencing and interruption scores that were higher than Dawadi et. al's [24] in measuring correlation of supervised task quality scoring with observed activity performance of dementia patients using supervised machine learning methods. Marcén et. al reported 78.86% accuracy of classifying AD and healthy controls through detecting nocturnal agitation [16], whereas a similar study revealed a much higher accuracy of 94.1% on detecting psycho-motor agitation [15]. The varieties in reported accuracy can be explained from different settings of the studies, including the choice of sensors, data collection methods and evaluation criteria. Machine learning algorithms with different cross-validation evaluation methods can yield different accuracy even if other settings are kept same. K-fold [3, 23] and leave-one-out cross-validation techniques [19] were commonly used in the reviewed studies. Kirste et. al [2] and Chikhaoui et. al [17] used leave-one-subject-out cross-validation whereas Riboni et. al [9] used leave-one-day-out cross-validation.

In this paper, we revisit the study performed by Kirste et. al [2] to detect AD from everyday motion data using accelerometers. We will use their data, test it with a new cross-validation method and introduce new dynamic and static interval based domain features and compare with their method of evaluation.

3 Dem@Care Study

3.1 Dataset

The study by Kirste et al. [2] attached an ankle-mounted accelerometer on each participant to collect their everyday motion behavior data. The raw data was the magnitude value of the normalized acceleration measured by the 3 axes accelerometer. The data set included motion data from 78 subjects (39 dyads). In the dataset, there was no data for 2 subjects; therefore, the number of subjects for this paper is 76 (38 dyads). An average of 53.4 hours of data were collected in the original study. We were shared with 5 types of data sets that were prepared by processing the data using a low pass filter with different frequencies (0.5, 5, 25, 50 and 250mHz). The data sets were named F1 to F5 corresponding to the five frequencies. For instance, the F1 data set was created from the raw data using a 0.5 mHz low pass filter.

3.2 Data Processing

As each subject were sampled at different times using the same sampling rate, the timestamps for each subject were slightly different even if we selected the same time window, for instance, between 22:00 on Day 1 and 22:00 on Day 2. Table 1 shows an example for the first two timestamps after 22:00 on Day 1 for two subjects. It can be seen that the data were sampled at different starting time after 22:00, one was at 22:02 and the other one was at 22:08. As all subsequent data were sampled in 50mHz (every 20 seconds) in this setting, the data for different subjects were therefore not sampled at the same time points. To make the samples comparable with each other, an interpolation of data is needed to synchronize data for all subjects. Ignoring this step may lead to undesirable results. To make the data points consistent with timestamps, we generated a reference array of timestamps between 22:00 on Day 1 to 22:00 on Day 2 based on a sampling rate used to generate the corresponding data set. The starting time for all 5 data sets was exact 22:00. Thus, an array containing reference timestamps can be obtained. For example, for the data set sampled in 50mHz, we obtained an array of 4320 timestamps (24 hour * 60 minutes * 60 seconds * 0.05). Then, we used MATLAB's *interp1* function [25] to generate the interpolated data. To prepare for interpolation, data of every subject were trimmed to 24-hour data. Only the data that has timestamps between 22:00 on Day 1 to 22:00 on Day 2 was kept. The rest of data were discarded in this paper.

Table 1: Example of Timestamps in Different Samples

Sample	**Motion Data**	
	Timestamp	***Value***
Participant 1	22:02	1.7084
	22:22	1.8173
Participant 2	22:08	2.1343
	22:28	2.0467

3.3 Model Settings

The machine learning model that gave the highest classification accuracy in the study by Kirste et al. [2] applied the model settings shown in Table 2. PCA was used to reduce the dimension of features generated by applying FFT on the original data. Six different time windows were tested in the original model. The 24-hour time window, which is from 22:00 on Day 1 to 22:00 on Day 2, was reported to give the best results. The better results on 24-hour time windows over other sub-windows of smaller duration may be because the time window includes a full cycle of a person's motion patterns from getting up from bed to going to sleep. This will cover different types of activities that a person would perform in a typical day. Therefore, we use the 24-hour window to build our models. They tested different machine learning classifiers and found that Quadratic Discriminant Analysis (QDA) classifier with the data set processed using 50mHz filter showed the best classification accuracy. To evaluate the classifiers, they adopted LOSOCV methodology.

Table 2: Model Settings for the Best Classification Accuracy

Category	**Setting**
Feature	4320 FFT
Reduced Features	5 PCA features
Time Window for Data Collection	22:00-22:00
Low Pass Filter on Datasets	50 mHz low pass filter
Machine Learning Classifier	Quadratic Discriminant Analysis

4 Revised Evaluation and Infusing Domain Knowledge

In the Dem@Care study, LOSOCV is employed for the evaluation of classifiers. This means that one of the person from the dyad will be present in the training set. This can bias the performance because both the person in the dyad lived together and may have done similar activities. To avoid such evaluation bias, we present LODOCV. In LODOCV, if there are N dyads present (or $2N$ total number of subjects), we train a classifier on $(N-1)$ dyads (or $(2N-2)$ subjects) and test on N^{th} dyad (or 2 subjects). We repeat the process N times and present the average performance. Performing LODOCV ensures that one dyad for testing will not be observed during training and biases due to similar actions of the subjects may be eliminated. To evaluate the effect of the new LODOCV evaluation on the Dem@Care dataset (discussed in Section 3.1), we extracted 4320 FFT features from the raw data and computed 5 PCA features on the entire feature data. Then, we trained different machine learning algorithms on the 5 data sets corresponding to different low pass filters (F1-F5). The machine learning classifiers used in this paper were Support Vector Machine (SVM), Random Forest (RF), and Quadratic Discriminant Analysis (QDA). For RF classifier, the number of trees was 100. LODOCV were performed with all machine learning classifiers along with LOSOCV to compare different evaluation methods.

4.1 Time and Frequency Features

In order to investigate the performance of revised evaluation methodology, we extract the following additional 10 time and frequency domain features [26] from the different accelerometer datasets within a time window of 24 hours:

- Mean, maximum, minimum, standard deviation, inter-quartile range,
- Total average power of power spectral density,
- Spectral entropy,
- DC component of energy,
- Normalized energy, and
- FFT entropy

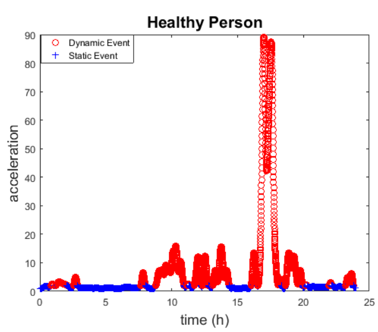
These features have been used earlier for activity recognition tasks. We compare and combine the time / frequency domain features with the proposed features, which are described in the next section.

4.2 Dynamic/Static Interval based Features

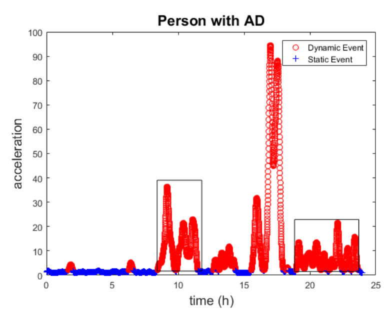
People with dementia / AD may exhibit different behaviours than healthy older adults, especially at different times of the day. This is referred to as 'sundowning' effect, where a person with dementia / AD can become more agitated, aggressive or confused during late afternoon or early evening [27]. Hsu et al. [21] show that gait and balance based quantitative measurements can be good indicators for early diagnosis of AD. We extend this idea to measure dynamic and static intervals from the accelerometer readings. A dynamic or static interval is defined as the total time period during which a person has more motor activities or not. The level of activity was deduced from the acceleration magnitude values that was consecutively higher or lower than a pre-defined value. This pre-defined threshold was empirically set to 4 in the unit of acceleration magnitude. The choice of the value 4 was based on the best accuracy obtained from multiple experiments using values range between $1 - 5$. This threshold will determine dynamic and static intervals in the overall motion activities of a person during the whole day. Figure 1a and 1b shows the dynamic intervals (in red color) and static intervals (in blue color) for a dyad that has a healthy person and another with AD at different times of the day. We can visually observe that the person with AD is more active during the later part of the day. We extracted 12 features based on these dynamic and static intervals. The features were extracted from the entire 24 hour data to obtain an overall motion pattern. Then, the day's data was divided into four 6-hour periods to capture motion information that may arise due to sundowning effect or other behaviours in specific quarters of the day. The following is the table of the extracted features:

Number of points in each dynamic and static intervals in 24 hours
Number of dynamic and static intervals in 24 hours
Number of dynamic and static intervals from 12:01 am - 6:00 am
Number of dynamic and static intervals from 6:01 am - 12:00 pm
Number of dynamic and static intervals from 12:01 pm - 6:00 pm
Number of dynamic and static intervals from 6:01 pm - 12:00 am

Table 3: The Extracted Domain Features



(a) Healthy Person.



(b) Person with AD.

Fig. 1: Dynamic and Static intervals of a dyad that comprises of a healthy person and with AD .

5 Results

To compare the performance of various machine learning algorithms on different versions of the Dem@Care datasets, we did the following experiments:

1. Compare the LOSOCV and LODOCV on the original 5 PCA features. This will act as a baseline for our analysis.

2. Compare the LOSOCV and LODOCV on time/frequency features, dynamic and static interval based features and their combinations with PCA features. This experiment will investigate the effect of new features in improving the baseline.

We used area under the ROC curve (AUC) as the performance metric, averaged across different folds.

5.1 PCA Features Baseline

Table 4 shows the AUC values on different datasets with 5 PCA features using LOSOCV and LODOCV evaluations. The number of PCA components is set to be the same as the work of Kirste et al. [2] for fair comparison. The best AUC obtained by LOSOCV is 0.73 for datasets F4 and F5 and 0.64 for LODOCV for dataset F4 using the QDA classifier. It is to be noted that in the LOSOCV evaluation 75 participants are used for training and for LODOCV 74 participants (or 37 dyads) were used for training. There is a difference of only 'one' participant during the training phase; however, the effect of performance is significant. This experiment highlights the bias that may be induced by including a person from the dyad to be tested in the training set as it may lead to identifying similar activities rather than patterns to discriminate AD and healthy persons. The drop in AUC by using LODOCV is expected as both the participants in the dyad were not observed during training. These results serve as a baseline for both types of evaluation methods when comparing the performance with other types of features.

Table 4: Performance of different classifiers and evaluation methods using 5PCA features.

Dataset	**Methods**	**LOSOCV**	**LODOCV**
F1	SVM	0.45	0.43
	RF	0.46	0.45
	QDA	0.56	0.51
F2	SVM	0.5	0.45
	RF	0.42	0.4
	QDA	0.54	0.49
F3	SVM	0.44	0.43
	RF	0.39	0.38
	QDA	**0.73**	**0.64**
F4	SVM	0.46	0.45
	RF	0.42	0.36
	QDA	**0.73**	0.58
F5	SVM	0.45	0.38
	RF	0.42	0.39
	QDA	0.55	0.5

5.2 Time, Frequency and Dynamic/Static Interval based features

The second experiment deals with evaluating different classifiers on the time, frequency and dynamic/static interval based features and their combinations with PCA features. We evaluated the classifiers on the following different feature regimes:

- 10 time and frequency features (10TF, discussed in Section 4.1).
- 12 new domain features (12D, discussed in Section 4.2).
- Combination of 10 time and frequency features and 5 PCA features (10TF+5PCA).
- Combination of 12 domain features and 5 PCA features (12D+5PCA).
- Combination of 12 domain features and 10 time and frequency features (12D+10TF).
- Combination of 12 domain features, 10 time and frequency features and 5 PCA features (12D+10TF+5PCA).

The classification results using 10TF, 10TF+5PCA and 12D features are shown in Table 5. The best AUC obtained for LOSOCV is 0.76 for F3 dataset using QDA classifier with combining time, frequency and PCA features, which is better than only using PCA features. This shows that adding additional time and frequency features improves the performance in the case of LOSOCV evaluation. For LODOCV, the best AUC is 0.71 achieved by using F3 dataset, RF classifier with domain features. The AUC outperforms the combinations of PCA, time and frequency features for the case of LODOCV evaluation. Therefore, this experiment suggests that domain specific features improve the performance of LODOCV evaluation.

To analyze the effect of domain features together with time, frequency and PCA features, we combine these features in different ways and compare their results in Table 6. With LOSOCV evaluation, the highest AUC of 0.74 is obtained for dataset F3 with QDA classifier using 12D+10TF+5PCA features. This AUC is less than the best AUC obtained by combining 10TF and 5PCA. This shows that dynamic/static interval based domain features do not help much for the LOSOCV evaluation. On the contrary, the best AUC of 0.73 is obtained using dataset F4, RF classifier, 12D+10TF+5PCA feature combination and LODOCV. This AUC is better than that of using domain features only (as shown in Table 5). This shows that combining domain features with time, frequency and PCA features improves the classification performance in patients with AD and healthy controls.

Through these experiments, we corrected the evaluation strategy of the Dem-@Care study, and improved the classification performance in detecting AD from everyday motion data by infusing domain knowledge features along with time, frequency and PCA features.

6 Conclusion

Detecting AD from everyday motion behaviour is a challenging problem. In this paper, we revisited the Dem@Care study that identifies people with AD by

Table 5: Performance of different classifiers and evaluation methods using 10TF, 10TF+5PCA and 12D features.

Feature	**Methods**	**LOSOCV**			**LODOCV**		
		10TF	**10TF+5PCA**	**12D**	**10TF**	**10TFD+5PCA**	**12D**
F1	SVM	0.48	0.48	0.52	0.49	0.49	0.51
	RF	0.52	0.45	0.46	0.46	0.46	0.56
	QDA	0.61	0.61	0.54	0.5	0.53	0.46
F2	SVM	0.45	0.45	0.52	0.45	0.45	0.55
	RF	0.55	0.54	0.41	0.48	0.42	0.62
	QDA	0.69	0.64	0.56	0.54	0.52	0.47
F3	SVM	0.52	0.52	0.52	0.52	0.52	0.67
	RF	0.58	0.48	0.36	0.49	0.45	**0.71**
	QDA	0.68	**0.76**	0.60	0.53	0.61	0.46
F4	SVM	0.53	0.53	0.42	0.52	0.52	0.65
	RF	0.42	0.44	0.37	0.35	0.40	0.64
	QDA	0.68	0.73	0.42	0.52	0.57	0.57
F5	SVM	0.44	0.44	0.49	0.38	0.45	0.53
	RF	0.4	0.4	0.46	0.34	0.40	0.58
	QDA	0.58	0.52	0.57	0.49	0.46	0.44

capturing their everyday motion data. Their evaluation method uses a subject from the dyad in training while the other is being tested, which can induce bias in the overall performance of the machine learning methods. To handle this problem, we proposed an evaluation methodology that keeps a full dyad out for testing without influencing the training set. Then, we propose new dynamic and static intervals features that are derived from different motor behaviours of persons with AD at different intervals of the day. By infusing domain features along with time, frequency features and the features used in Dem@Care study, significant improvement in performance was obtained using LODOCV evaluation strategy. In future, we will work on automating the threshold to distinguish static and dynamic intervals and extract more discriminatory features.

References

1. Association, A.: 2017 Alzheimer's disease facts and figures. Alzheimer's & Dementia: The Journal of the Alzheimer's Association **13**(4) (2017) 325–373
2. Kirste, T., Hoffmeyer, A., Koldrack, P., Bauer, A., Schubert, S., Schröder, S., Teipel, S.: Detecting the effect of Alzheimer's disease on everyday motion behavior. Journal of Alzheimer's Disease **38**(1) (2014) 121–132
3. Janjua, Z.H., Riboni, D., Bettini, C.: Towards automatic induction of abnormal behavioral patterns for recognizing mild cognitive impairment. In: Proceedings of the 31st Annual ACM Symposium on Applied Computing, ACM (2016) 143–148
4. Karakostas, A., Meditskos, G., Stavropoulos, T.G., Kompatsiaris, I., Tsolaki, M.: A Sensor-Based Framework to Support Clinicians in Dementia Assessment: The

Table 6: Results for Domain Features Combined with Statistical and Spectral Features

Dataset	Method	LOSOCV			LODOCV		
		12D+5PCA	12D+ 10TF	12D+10TF+5PCA	12D+5PCA	12D+10TF	12D+10TF+5PCA
		AUC	AUC	AUC	AUC	AUC	AUC
F1	SVM	0.41	0.48	0.48	0.51	0.51	0.51
	RF	0.37	0.49	0.49	0.61	0.55	0.58
	QDA	0.58	0.58	0.59	0.46	0.51	0.49
F2	SVM	0.62	0.45	0.45	0.55	0.55	0.55
	RF	0.52	0.52	0.53	0.54	0.57	0.48
	QDA	0.54	0.62	0.57	0.49	0.44	0.48
F3	SVM	0.38	0.52	0.52	0.67	0.48	0.48
	RF	0.38	0.45	0.47	0.64	0.59	0.58
	QDA	0.67	0.69	**0.74**	0.41	0.39	0.37
F4	SVM	0.38	0.53	0.53	0.65	0.48	0.48
	RF	0.36	0.35	0.35	0.72	0.69	**0.73**
	QDA	0.46	0.54	0.56	0.58	0.53	0.52
F5	SVM	0.54	0.44	0.44	0.53	0.55	0.55
	RF	0.45	0.49	0.46	0.54	0.59	0.56
	QDA	0.55	0.60	0.58	0.47	0.46	0.49

Results of a Pilot Study. Advances in Intelligent Systems and Computing **376** (2015) 213–221
5. Dem@Care: Dem@Care Project. `http://www.demcare.eu/` Accessed on 2017-08-13.
6. Tung, J., Semple, J., Woo, W., Hsu, W.: Ambulatory Assessment of Lifestyle Factors for Alzheimer's Disease and Related Dementias. AAAI Spring Symposium (2011) 50–54
7. Ando, B., Baglio, S., Lombardo, C.O., Marletta, V.: A multisensor data-fusion approach for ADL and fall classification. IEEE Transactions on Instrumentation and Measurement **65**(9) (2016) 1960–1967
8. Akl, A., Taati, B., Mihailidis, A.: Autonomous unobtrusive detection of mild cognitive impairment in older adults. IEEE Transactions on Biomedical Engineering **62**(5) (2015) 1383–1394
9. Riboni, D., Bettini, C., Civitarese, G., Janjua, Z.H., Helaoui, R.: SmartFABER: Recognizing fine-grained abnormal behaviors for early detection of mild cognitive impairment. Artificial Intelligence in Medicine **67** (2016) 57–74
10. Dawadi, P.N., Cook, D.J., Schmitter-Edgecombe, M.: Automated Cognitive Health Assessment Using Smart Home Monitoring of Complex Tasks. IEEE Trans. on Systems, Man, and Cybernetics – Part C: Applications and Reviews **43**(6) (2013) 1302–1313
11. Arifoglu, D., Bouchachia, A.: Activity Recognition and Abnormal Behaviour Detection with Recurrent Neural Networks. Procedia Computer Science **110** (2017) 86–93
12. Sacco, G., Joumier, V.V., Darmon, N., Dechamps, A., Derreumaux, A., Lee, J.H.H., Piano, J., Bordone, N., Konig, A., Teboul, B., David, R., Guerin, O., Bremond, F.F., Robert, P.: Detection of activities of daily living impairment in

Alzheimer's disease and mild cognitive impairment using information and communication technology. Clinical Interventions in Aging **7** (2012) 539–549
13. Crispim-junior, C.F., Joumier, V., Hsu, Y.l., Pai, M.c., Chung, P.c., Dechamps, A., Robert, P., Bremond, F.: Alzheimer's patient activity assessment using different sensors. Gerontechnology **11**(2) (2012) 266–267
14. Nagels, G., Engelborghs, S., Vloeberghs, E., Van Dam, D., Pickut, B.a., De Deyn, P.P.: Actigraphic measurement of agitated behaviour in dementia. International Journal of Geriatric Psychiatry **21**(4) (2006) 388–393
15. Pedro, S., Quintas, J., Menezes, P.: Sensor-based detection of Alzheimer's disease-related behaviors. In: The International Conference on Health Informatics, IFMBE Proceedings. (2014)
16. Marcén, A.C., Carro, J., Monasterio, V.: Wearable Monitoring for the Detection of Nocturnal Agitation in Dementia. Proceedings of the 6th International Joint Conference on Pervasive and Embedded Computing and Communication Systems (January) (2016) 63–69
17. Chikhaoui, B., Ye, B., Mihailidis, A.: Ensemble Learning-Based Algorithms for Aggressive and Agitated Behavior Recognition. In: Proc. of the 10th International Conference on Ubiquitous Computing & Ambient Intelligence (UCAMI 2016). (2016)
18. David, R., Rivet, A., Robert, P.H., Mailland, V., Friedman, L., Zeitzer, J.M., Yesavage, J.: Ambulatory actigraphy correlates with apathy in mild Alzheimer's disease. Dementia **9**(4) (2010) 509–516
19. Miranda, D., Favela, J., Ibarra, C.: Detecting State Anxiety when Caring for People with Dementia. Methods Inf Med **56** (2017) 55–62
20. Kuhlmei, a., Walther, B., Becker, T., Müller, U., Nikolaus, T.: Actigraphic daytime activity is reduced in patients with cognitive impairment and apathy. European Psychiatry **28**(2) (2013) 94–97
21. Hsu, Y.L., Chung, P.C., Wang, W.H., Pai, M.C., Wang, C.Y., Lin, C.W., Wu, H.L., Wang, J.S.: Gait and Balance Analysis for Patients With Alzheimer's Disease Using an Inertial-Sensor-Based Wearable Instrument. IEEE Journal of Biomedical and Health Informatics **18**(6) (2014)
22. Richter, J., Wiede, C., Hirtz, G.: Mobility assessment of demented people using pose estimation and movement detection an experimental study in the field of ambient assisted living. ICPRAM 2015 - 4th International Conference on Pattern Recognition Applications and Methods, Proceedings **2** (2015) 22–29
23. Alam, M.A.U., Roy, N., Holmes, S., Gangopadhyay, A., Galik, E.: Automated Functional and Behavioral Health Assessment of Older Adults with Dementia. First IEEE Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), Washington D.C. June 2016 (May) (2016)
24. Dawadi, P.N., Cook, D.J., Schmitter-Edgecombe, M.: Automated Cognitive Health Assessment from Smart Home-Based Behavior Data. IEEE Journal of Biomedical and Health Informatics **20**(4) (2016) 1188–1194
25. The MathWorks, I.: MATLAB R2016b
26. Khan, S.S., Karg, M.E., Kuli, D., Hoey, J.: Detecting falls with x-factor hidden markov models. Applied Soft Computing **55** (2017) 168 – 177
27. (GB), A.S.: Sundowning – behaviour changes. `https://www.alzheimers.org.uk/info/20064/symptoms/87/behaviour_changes/8` Accessed on 2018-01-11.