proceedings

MDPI

Proceedings

A Convolutional Neural Network-Based Method for Human Movement Patterns Classification in Alzheimer's Disease †

Santos Bringas [1,*], Sergio Salomón [2], Rafael Duque [3], José Luis Montaña [3] and Carmen Lage [1]

1 Cognitive Disorders Unit, Department of Neurology, Marqués de Valdecilla University Hospital (HUMV) Valdecilla Biomedical Research Institute (IDIVAL), 39008 Santander, Spain; clage@idival.org

2 Axpe Consulting Cantabria SL, 39600 Camargo, Spain; ssalomong@axpecantabria.com

3 Department of Mathematics, Statistics and Computer Science, Universidad de Cantabria, 39005 Santander, Spain; rafael.duque@unican.es (R.D.); joseluis.montana@unican.es (J.L.M.)

* Correspondence: sbringas@idival.org (S.B.)

† Presented at the 13th International Conference on Ubiquitous Computing and Ambient Intelligence UCAmI 2019, Toledo, Spain, Spain, 2–5 December 2019.

Published: 21 November 2019

Abstract: Alzheimer's disease (AD) constitutes a neurodegenerative pathology that presents mobility disorders as one of its earliest symptoms. Current smartphones integrate accelerometers that can be used to collect mobility data of Alzheimer's patients. This paper describes a method that processes these accelerometer data and a convolutional neural network (CNN) that classifies the stage of the disease according to the mobility patterns of the patient. The method is applied in a case study with 35 Alzheimer's patients, in which a classification success rate of 91% was obtained.

Keywords: mobile and ubiquitous health; Alzheimer; convolutional neural networks

1. Introduction

The introduction of the paradigm of ubiquitous computing [1] was a revolution in the way of interacting with computers. According to this paradigm, the software systems include features that automatically recognize the user requirements and put mechanisms in actions to satisfy them. For this purpose, sensors and small technological devices are distributed in the environment to recognize the human activity and to adapt the software mechanisms to the user requirement in each moment.

Healthcare is an area of growing interest in which to apply the paradigm of interaction proposed by ubiquitous computing [2]. This is a consequence of the capability of the small devices that include sensors to capture information of interest for diagnosing pathologies (physical activity, heart rate, mobility patterns, falls, etc.) with any interaction effort from the patient. This situation involves a new challenge from the computational point of view, such as the design of mechanisms that process an important amount of data from which to generate new information that enables health professionals to monitor and diagnose pathologies.

Alzheimer's disease manifests mobility disorders as one of its earliest symptoms [3]. Current smartphones integrate several sensors (GPS, accelerometer, gyroscope, etc.) that capture information related to user mobility. Therefore, it is worth considering the exploitation of these smartphone sensors to assess the mobility problems associated with Alzheimer's disease [4]. In this way, the patient benefits from the interaction advantages of the ubiquitous computing paradigm and obtains constant monitoring of their mobility patterns without going to a specialized laboratory. To address this challenge, this work proposes a methodology that compresses the size of the data collected by accelerometers of patients with Alzheimer's disease and uses convolutional neural

networks(CNNs) to predict the stage of the pathology. The methodology is applied to a case study in which 35 patients participate.

The article is structured in five additional sections. Section 2 reviews the main scientific contributions in the field of ubiquitous technologies to monitor and assess Alzheimer's disease. Section 3 presents our methodology to study the stage of Alzheimer's disease using mobility data. Section 4 describes a case study in which the methodology is applied. Section 5 collects the conclusions drawn from this work.

2. Related Work

Alzheimer's disease (AD) constitutes a neurodegenerative disease that leads to progressive cognitive decline and disability, with severe familiar and socio-economic related consequences. As the first cause of dementia, in 2015 it was estimated that almost 47 million people worldwide were affected by dementia, and, due to the progressive aging of our societies, these figures are expected to reach 131 million by 2050 [5], which has led the World Health Organization to conclude that AD and other dementias should be considered as a global public health priority [6]. From the clinical perspective, it is important to track the progression of the disease for optimal treatment management and prevention of complications. In this regard, the severity of the disease is evaluated in clinical practice through standardized interviewing tools, such as the Global Deterioration Scale [7]. This scale establishes different stages: no objective cognitive impairment (GDS 1-2); mild cognitive impairment (MCI) when there is no significant functional impact (GDS 3); and dementia when the cognitive impairment results in some functional impairment (mild dementia or GDS 4, moderate dementia or GDS 5, moderately severe dementia, or GDS 6 and severe dementia or GDS 7).

Over the last decade, technology, and more specifically sensor-based wearable devices, have emerged as a novel resource with broad applications in the field of cognitive disorders. Since currently there are no available curative treatments for AD, a main focus of research are the improvements in quality of life and safety aids, and sensor-based systems have been proposed for many monitoring purposes in patients with dementia, as the evaluation of sleep quality [8,9], the detection of leaving-bed episodes [10] or the estimation of falls risk [11,12], which offers the opportunity to prevent accidents or to evaluate the response to treatment adjustments. Furthermore, sensor-based systems can also yield a great amount of information that can be applied to diagnostic purposes. A key advantage in contrast to conventional approaches is that portable devices can provide a source of ecological information and, therefore, certain kinds of data that cannot be obtained in the clinical setting or supervised laboratory environments. For instance, different works report the methodology for detecting gait episodes and estimating gait parameters from accelerometric signals [13,14], which have been demonstrated to be able to distinguish both patients with dementia [14] and even MCI [15] from healthy controls. Additionally, there is increasing evidence that some gait abnormalities are present since very early stages of dementia, even when they are still unapparent in clinical assessment. Thus, Gillain et al. [16] described that MCI patients who developed dementia in the subsequent three years showed lower gait speed, symmetry, and regularity measured by a tri-axial accelerometer than MCI patients who remained stable, which suggests that gait abnormalities can constitute a preclinical biomarker and that an approach based on accelerometry is technically feasible. However, beyond intrinsic characteristics of gait, wearable devices also can provide insight into daily activity patterns. Actigraphy has proved to be useful for the monitoring of behavioral changes, such as apathy, probably the most frequent neuropsychiatric symptom of AD [17–19]. But the assessment of activity patterns can also be of diagnostic value. Interestingly, Watts et al. [20] found significant differences in the intra-individual variability of physical activity between mild AD patients and cognitively intact older adults. Similarly, Kirste et al. [21] studied everyday motion behavior by means of tri-axial accelerometers in 23 dyads (23 AD patients and their cognitively healthy partners) and obtained a classification accuracy of 91% for the discrimination between AD patients and healthy controls. Additionally, motion features were significantly correlated with scores in the Mini-Mental State Examination [22], probably the commonest

brief cognitive test and mostly used to track the progression of AD, which could imply a relationship between daily motion behavior and the stage of the disease. Previous works have been published focused on the staging of AD or other dementias employing sensor-based systems [23,24]. However, these works did not explore the potential of the CNNs to classify the stage of AD.

3. Method

We address the problem of Alzheimer's staging classification using patients' accelerometers data. The time-series data provided by accelerometers require some preprocessing steps to analyze them effectively since it may contain noise, have heterogeneous lengths or may not be time regular. Once data is preprocessed, it can be exploited by a machine learning algorithm to obtain a model that classifies new data. In this methodology, we apply models based on deep learning techniques using Convolutional Neural Networks.

3.1. Problem Description

The addressed problem in this work is the classification of accelerometer data sequences according to the different Alzheimer's disease stages. The classes defined by these stages are early (corresponding to GDS 2 and 3), middle (GDS 4 and 5), and late (GDS 6 and 7). Then, a classifier should use the patient records of accelerometer movement to predict one of the three categories.

Since the input data of the problem comes from accelerometer sensors, there will be at least one accelerometer sequence for each patient in the dataset. These sequences, data streams registered over time, will contain values for the acceleration changes in the axes X, Y, and Z, and those will be the three data features to be considered in our methodology along with the temporal dimension.

3.2. Data Preprocessing

In the preprocessing phase, we aim to transform all data sequences to the same length and homogenize periods between data points. The first step is the data homogenization according to the temporal dimension; i.e., a transformation of the time intervals between consecutive points to be equal. The data captured by a device sometimes can have very high resolution (i.e., points with very short periods) and result in too long sequences. Besides, these high-frequency points may be repeating very similar values that do not provide useful information for the classification task. To avoid this situation and also reduce considerably the dataset size, we set a time resolution parameter and group points by this value (e.g., every 0.5 s), keeping the average value of a set of points in the same fragment to represent them. This step will work as lossy data compression, removing redundant information, and will avoid data unbalancing issues.

Next, the data is divided into several segments in order to reduce sequence lengths and increase the total number of samples. The number of segments to partition each sequence should be chosen depending on the specific dataset. This step serves as a data augmentation of the samples with which the model will be trained and tested. As a result of the preprocessing, each data sequence of a patient is transformed in a set of shorter fixed-length sequences with homogeneous periods, hence distributing points uniformly in time.

3.3. Convolutional Neural Network Classifier

Among the several machine learning models that can be used for the classification task, we focus our work on the recent deep learning developments. Because of the characteristics of this problem, we expect the movement data to exhibit complex patterns sequences of mixed length over time. Thus, deep learning methods present as suitable classifiers for this data since they can take advantage of the internal structure of the sequences. Moreover, CNNs, models of the deep learning family with growing popularity, can recognize hierarchies of patterns from smaller ones without increasing the model complexity.

The CNN consists of an artificial neural network structure with input, output, and several hidden layers. These hidden layers are based on the *convolution*, a linear operation between the input and a kernel (or filter) that serves as a feature detector. As CNNs work internally with convolutions in different sliding windows, these models can identify patterns locally and, therefore, make a better distinction of which are representative of each class.

The overview of our architecture for the neural network is shown in Figure 1. In this proposal, three 1D Convolutional layers have been used, each one including a Batch Normalization (BN) layer and an Average Pooling (Avg Pool) layer. Each convolutional layer applies linear rectification (Rectified Linear Unit or ReLU) as its activation function. This set of layers allows the model to extract patterns from the input data (commonly known as feature learning). After this, the output of the last convolution is flattened and used as input for the last two layers: a fully connected (FC) layer and a *Softmax* layer with three neurons. This portion of the network is responsible for the classification task by computing the probability that the input given belongs to a certain label. This kind of multi-layer structure, with Batch Normalization, pooling, and increasing filter sizes, has proven to be effective for time-series analysis [25,26].

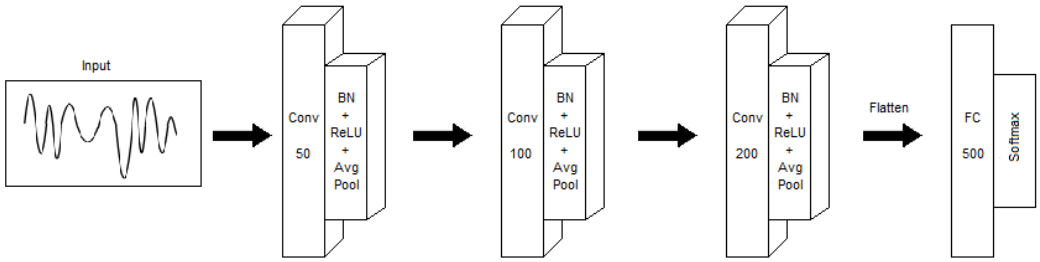

Figure 1. Architecture used for the Convolutional Neural Network (CNN). BN stands for Batch Normalization, ReLU for Rectified Linear Unit, and FC for fully connected.

In our model, we use 1-Dimensional (or 1D) Convolutional layers. The 1D CNNs are well-fitting to manage time-series, like accelerometer or gyroscope data, where fixed-length data segments are analyzed sequentially to detect local patterns. Since we work with accelerometer data, the input features (or channels) for the network will be the variables X, Y, and Z; then, 1D convolution will process each channel independently along the time progression of the sequence. The output of a convolutional layer is called *feature map* and represents the detection of different patterns from the input.

The Batch Normalization method normalizes the output of each activation so that changes in weights of these layers do not affect severely the next layers [27]. This technique allows faster convergence in the training process and makes the model more robust to different initializations. The average pooling method applies dimensionality reduction (called *downsampling*) to the feature map by averaging a group of values inside a window. This operation reduces the complexity of the output in each convolutional layer and avoids overfitting.

Last, two fully-connected layers are placed to obtain the model prediction. These two layers learn non-linear combinations from the input (the flatten feature maps) to identify the representative patterns for each class. Then, the label is chosen by the Softmax classification function. A dropout layer has been placed before the fully-connected part to avoid redundant neurons activating for the same feature (co-adaptation) [28], thereby preventing model overfitting.

When working with deep neural networks, the number of the hyperparameter combinations is huge due to its complex multi-layer design: number of layers, layer size, number of filters, number of fully-connected neurons, etc. Different variations of those in the network model should be tested with the specific data to select the parameters for the ideal configuration.

4. Results

We applied the methodology described in Section 3 in a case study with a specific dataset to analyze the performance of our model. After preprocessing the data and training the CNN model, we compared the results obtained from the prediction against other common supervised learning classifiers. The method implementation and the experimentation, carried out under the programming language Python, was supported by the *Keras* library to create and train the CNN model, and the libraries *NumPy* and *Scikit-Learn* to manipulate and preprocess data.

4.1. Data Description

The dataset from the case study was introduced in [23]. In this experiment, data were collected from 35 patients with Alzheimer's disease carrying a mobile device in a small pocket placed and oriented by a neuropsychologist. The patients were labeled as follows: seven patients in early-stage, 18 patients in middle-stage, and 10 patients in late-stage.

The Android device accelerometers took measures with a sampling rate of 8 Hz in a range of ± 3.28 g. The data are affected by the different activities carried out by the patients during the recordings. In addition, small errors in the placement of the devices by the neuropsychologist are contemplated as part of the study. The patient movements were recorded for around a week in short sessions each day, resulting in a total of 187 sequences. The longest accelerometer sample contains more than a half million points, but most of the sequences have shorter lengths. In Figure 2, two samples from different patients are represented in each axis.

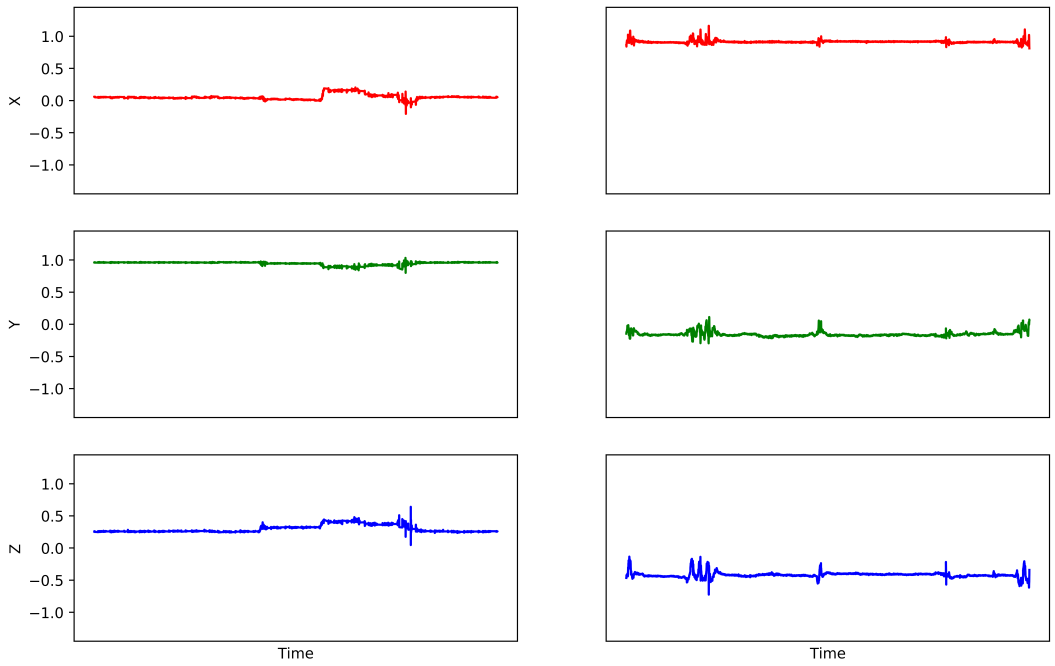

Figure 2. Accelerometer data sequences for two different patients (early and middle stage, respectively).

After applying the time homogenization of our preprocessing methodology with a time resolution parameter of 0.1 s, we obtained data sequences with 1.5 h of maximum length (around 54k data points). Then, we applied the data partitioning step and divided all sequences into five segments. This process gave us 935 data traces with a maximum sequence length of 18 min (approximately 11k data points). Because of this process, we considerably increased the number of samples that could be used in the learning process.

4.2. Model Training

In order to evaluate the proposed model, the dataset was divided into training and testing sets: 80% of the data were considered for training and 20% of the data for testing. Because of the data augmentation process performed in the previous preprocessing (where each original sequence was divided into five segments), we considerably increased the number of samples of each patient and each class. This increments the probability of two samples from the same patient, or even from the same original sequence, to be assigned into the different sets (training or testing). Thus, different characteristics can be learned from all patients and the training dataset is not unbalanced with data from a single patient or a single class. Therefore, we performed a random data selection. Finally, the result of the train/test split gave us a total of 748 sequences to train and 187 to test (44 early-stage, 95 middle-stage, and 48 late-stage).

For the training of our network model, we chose the optimal setting of hyperparameters for the dataset used. After several tests, we selected the Adam optimization algorithm [29] with a learning rate of 0.001; the loss function used is categorical cross-entropy, and the model was trained for 300 dataset iterations (*epochs*). Here, the model that obtained the lowest training loss was kept.

4.3. Model Evaluation

After selecting the optimal model configuration for our data in the training process, we assessed the model prediction. In addition, we compared these results against a set of common supervised learning classifiers: Support-Vector Machines (SVMs), Random Trees (RTs) , Random Forest (RF), and Multilayer Perceptron (MLP) (with logistic activation). These models were originally compared in [23], in which the MLP was the most successful classifier.

Table 1 shows the results obtained with the different models used for classifying the dataset. Here, we can see the percentage of classification success for each label (early, middle, and late) obtained by each of the trained models. It should be noted that, in the case of CNN, the number of samples was multiplied fivefold as a result of the data augmentation step in our preprocessing method and, therefore, it had more examples in its training and validation (748 training samples and 187 testing samples). The rest of the models used 150 samples for training (80 % of the original dataset) and 37 for testing (20 % of the original data).

Table 1. Classification results the given dataset. CNN stands for Convolutonal Neural Network, MLP stands for Multi-Layer Perceptron, RT for Random Trees, RF for Random Forest, and SVM for Support-Vector Machines. The best results are highlighted in bold.

Technique	Early-Stage. .	Middle-Stage. .	Late-Stage. .	Total
CNN	. . . . 89%. .	. . 93%	. . **91**%	. . . **91**%
MLP	. . . . **100**%. .	. . **100**%	. . 50%	. . . 83%
RT	. . . . 0%. .	. . 100%	. . 0%	. . . 50%
RF	. . . . 0%. .	. . 100%	. . 0%	. . . 50%
SVM	. . . . 0%. .	. . 100%	. . 0%	. . . 50%

In light of the results obtained, we compared the tested models by its success rate for each class samples and the total samples. CNN is the model that presents the highest total success rate (91%) and maintains high similar rates for the three classes (89%, 93%, and 91%). Thus, it seems to be the best model to distinguish all stages. The MLP model exhibit higher success rates for early and middle stages (100%) than CNN, but it had worse performance recognizing the late stage. Finally, the other models (RT, RF, and SVM) do not identify early and late stages correctly (only succeeding in the middle stage recognition). In conclusion, we see CNN as a promising model for the stage classification task, which improves the performance of other neural network models like the MLP.

5. Conclusions

Alzheimer's disease is a pathology with a growing prevalence today. One of its first symptoms is the appearance of mobility disorders. This work has described a method to process mobility data captured by smartphone sensors and to establish a relationship with the stage of the disease by means of Convolutional Neural Networks.

The method was applied in a case study with 35 patients. The results obtained a 91% success in classifying the stage of the disease (as early, middle, or late). Thus, the success rate of the model can be considered as promising. Currently, we propose to exploit this method within a more global technological environment that allows health personnel to manage different sources of information on the patient's evolution. This new objective will involve analyzing information on other symptoms of Alzheimer's disease (loss of recent memories, language problems, etc.).

Author Contributions: Conceptualization, S.B., R.D. and J.L.M.; methodology, S.B., R.D. and J.L.M.; software, S.B.; validation, S.B., S.S. and R.D.; formal analysis, S.B.; investigation, S.B., S.S. and C.L.; data curation, S.B. and S.S.; writing—original draft preparation, S.B., R.D. and C.L.; writing—review and editing, S.S., R.D. and J.L.M.; visualization, S.S.; supervision, R.D. and J.L.M.

Funding: The authors want to acknowledge the financial support from FEDER (Fondo Europeo de Desarrollo Regional) and The Government of Cantabria for the project 2018/INN/9 within the program INNOVA 2018, and the ISCIII (Instituto de Salud Carlos III) and Ministerio de Economía y Competitividad, Gobierno de España for the project PI17/00936.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AD	Alzheimer's Disease
CNN	Convolutional Neural Network
GDS	Global Deterioration Scale
MCI	Mild Cognitive Impairment
MLP	Multi-Layer Perceptron
RF	Random Forest
RT	Random Tree
SVM	Support-Vector Machine

References

1. Weiser, M.; Gold, R.; Brown, J.S. The Origins of Ubiquitous Computing Research at PARC in the Late . 1980s. *IBM Syst. J.* **1999**, *38*, 693–696, doi:10.1147/sj.384.0693.
2. Deen, M.J. Information and Communications Technologies for Elderly Ubiquitous . Healthcare in a Smart Home. *Pers. Ubiquitous Comput.* **2015**, *19*, 573–599, doi:10.1007/s00779-015-0856-x.
3. R Varma, V.; Watts, A. Daily Physical Activity Patterns During the Early Stage of . Alzheimer's Disease. *J. Alzheimer's Dis. JAD* **2016**, *55*, doi:10.3233/JAD-160582.
4. Kourtis, L.C.; Regele, O.B.; Wright, J.M.; Jones, G.B. Digital biomarkers for Alzheimer's disease: The mobile/wearable . devices opportunity. *Npj Digit. Med.* **2019**, *2*, 9, doi:10.1038/s41746-019-0084-2.
5. Prince, M.; Wimo, A.; Guerchet, M.; Ali, G.; Wu, Y.; Prina, M. *World Alzheimer Report 2015*; Alzheimer's Disease International: London, UK, 2015; pp. 1–92, doi:10.1111/j.0963-7214.2004.00293.x.
6. World Health Organization. Dementia: A Public Health Priority. World Health Organization: Geneva, Switzerland, 2012; pp. 1–4, ISBN 9789241564458.
7. Reisberg, B.; Ferris, S.H.; de Leon, M.J.; Crook, T. The Global Deterioration Scale for assessment of primary degenerative . dementia. . *Am. J. Psychiatry* **1982**, *. 139*, 1136–1139, doi:10.1176/ajp.139.9.1136.
8. Leger, D.; Elbaz, M.; Dubois, A.; Rio, S.; Mezghiche, H.; Carita, P.; . Stemmelin, J.; Strauss, M. Alzheimer's Disease Severity is Not Significantly Associated with . Short Sleep: Survey by Actigraphy on 208 Mild and Moderate Alzheimer's . Disease Patients. *J. Alzheimer's Dis. JAD* **2017**, *. 55*, 321–331, doi:10.3233/JAD-160754.

9. Camargos, E.F.; Louzada, F.M.; Nobrega, O.T. Wrist actigraphy for measuring sleep in intervention studies with . Alzheimer's disease patients: Application, usefulness, and challenges. *Sleep Med. Rev.* **2013**, *17*, 475–488, doi:10.1016/j.smrv.2013.01.006.
10. Higami, Y.; Yamakawa, M.; Shigenobu, K.; Kamide, K.; Makimoto, K. High frequency of getting out of bed in patients with Alzheimer's . disease monitored by non-wearable actigraphy. *Geriatr. Gerontol. Int.* **2019**, *. 19*, 130–134, doi:10.1111/ggi.13565.
11. Gietzelt, M.; Feldwieser, F.; Gövercin, M.; Steinhagen-Thiessen, E.; . Marschollek, M. A prospective field study for sensor-based identification of fall . risk in older people with dementia. *Inform. Health Soc. Care* **2014**, *. 39*, 249–261, doi:10.3109/17538157.2014.931851.
12. Di Rosa, M.; Hausdorff, J.M.; Stara, V.; Rossi, L.; Glynn, L.; Casey, M.; . Burkard, S.; Cherubini, A. Concurrent validation of an index to estimate fall risk in community . dwelling seniors through a wireless sensor insole system: A pilot study. *Gait Posture* **2017**, *55*, 6–11, doi:10.1016/j.gaitpost.2017.03.037.
13. Hsu, Y.L.; Chung, P.C.J.; 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 J. Biomed. Health Inform.* **2014**, *. 18*, 1822–1830, doi:10.1109/JBHI.2014.2325413.
14. Gietzelt, M.; Wolf, K.H.; Kohlmann, M.; Marschollek, M.; Haux, R. Measurement of Accelerometry-based Gait Parameters in People with and . without Dementia in the Field. *Methods Inf. Med.* **2013**, *. 52*, 319–325, doi:10.3414/me12-02-0009.
15. Hausdorff, J.M.; Hillel, I.; Shustak, S.; Del Din, S.; Bekkers, E.M.J.; . Pelosin, E.; Nieuwhof, F.; Rochester, L.; Mirelman, A. Everyday Stepping Quantity and Quality Among Older Adult Fallers With . and Without Mild Cognitive Impairment: Initial Evidence for New Motor Markers . of Cognitive Deficits? *. J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **2018**, *73*, 1078–1082, doi:10.1093/gerona/glx187.
16. Gillain, S.; Drame, M.; Lekeu, F.; Wojtasik, V.; Ricour, C.; Croisier, J.L.; . Salmon, E.; Petermans, J. Gait speed or gait variability, which one to use as a marker of risk . to develop Alzheimer disease? A pilot study. *Aging Clin. Exp. Res.* **2016**, *. 28*, 249–255, doi:10.1007/s40520-015-0392-6.
17. David, R.; Mulin, E.; Friedman, L.; Le Duff, F.; Cygankiewicz, E.; Deschaux, . O.; Garcia, R.; Yesavage, J.A.; Robert, P.H.; Zeitzer, J.M. Decreased daytime motor activity associated with apathy in Alzheimer . disease: An actigraphic study. *. Am. J. Geriatr. Psychiatry Off. J. Am. Assoc. Geriatr. Psychiatry* **2012**, *. 20*, 806–814, doi:10.1097/JGP.0b013e31823038af.
18. Kuhlmei, A.; Walther, B.; Becker, T.; Muller, U.; Nikolaus, T. Actigraphic daytime activity is reduced in patients with cognitive . impairment and apathy. *Eur. Psychiatry J. Assoc. Eur. Psychiatr.* **2013**, *28*, 94–97, doi:10.1016/j.eurpsy.2011.04.006.
19. Zeitzer, J.M.; David, R.; Friedman, L.; Mulin, E.; Garcia, R.; Wang, J.; . Yesavage, J.A.; Robert, P.H.; Shannon, W. Phenotyping apathy in individuals with Alzheimer disease using . functional principal component analysis. *. Am. J. Geriatr. Psychiatry Off. J. Am. Assoc. Geriatr. Psychiatry* **2013**, *. 21*, 391–397, doi:10.1016/j.jagp.2012.12.012.
20. Watts, A.; Walters, R.W.; Hoffman, L.; Templin, J. Intra-Individual Variability of Physical Activity in Older Adults . With and Without Mild Alzheimer's Disease. *PLoS ONE* **2016**, *11*, e0153898, doi:10.1371/journal.pone.0153898.
21. 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. *J. Alzheimer's Dis.* **2014**, *. 38*, 121–132, doi:10.3233/JAD-130272.
22. Folstein, M.F.; Folstein, S.E.; McHugh, P.R. "Mini-mental state". A practical method for grading the cognitive . state of patients for the clinician. *J. Psychiatr. Res.* **1975**, *. 12*, 189–198.
23. Nieto-Reyes, A.; Duque, R.; Montaña, J.L.; Lage, C. Classification of Alzheimer's Patients through Ubiquitous . Computing. *Sensors* **2017**, *17*, 1679.
24. Duque, R.; Nieto-Reyes, A.; Martínez, C.; Montaña, J.L. Detecting Human Movement Patterns Through Data Provided by . Accelerometers. A Case Study Regarding Alzheimer's Disease. . In Proceedings of the. Ubiquitous Computing and Ambient Intelligence - 10th International . Conference (UCAmI 2016), San Bartolomé de Tirajana, Gran Canaria, Spain, 29 November–2 December. 2016; pp. 56–66, doi:10.1007/978-3-319-48746-5_6.
25. Ann Ronao, C.; Cho, S.B. Human activity recognition with smartphone sensors using deep . learning neural networks. *Expert Syst. Appl.* **2016**, *59*, doi:10.1016/j.eswa.2016.04.032.

26. Wang, Z.; Yan, W.; Oates, T. Time series classification from scratch with deep neural networks: A strong baseline. . In Proceedings of the. 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May. 2017; pp. 1578–1585, doi:10.1109/IJCNN.2017.7966039.
27. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing . Internal Covariate Shift. . In. Proceedings of the 32nd International Conference on International . Conference on Machine Learning (ICML'15), Lille, France, 7–9 July 2015; Volume 37, pp. . 448–456.
28. Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Improving neural networks by preventing co-adaptation of feature . detectors. *arXiv preprint* **2012**, arXiv:1207.0580.
29. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In. Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015), San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings.