Computer Methods and Programs in Biomedicine 220 (2022) 106841

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

A self-driven approach for multi-class discrimination in Alzheimer's disease based on wearable EEG

Eduardo Perez-Valero[a,c], Miguel Ángel Lopez-Gordo[b,c,*], Christian Morillas Gutiérrez[a,c], Ismael Carrera-Muñoz[d], Rosa M. Vílchez-Carrillo[d]

[a] Department of Computer Architecture and Technology, University of Granada, Spain
[b] Department of Signal Theory, Telematics and Communications, University of Granada, Spain
[c] Brain-Computer Interfaces Laboratory, Research Centre for Information and Communications Technologies, University of Granada, Spain
[d] Cognitive Neurology Group, Neurology Unit, Hospital Universitario Virgen de las Nieves, Granada, Spain

ARTICLE INFO

Article history:
Received 15 February 2022
Revised 25 March 2022
Accepted 25 April 2022

Keywords:
Alzheimer's disease
automated detection
EEG
machine learning

ABSTRACT

Early detection is critical to control Alzheimer's disease (AD) progression and postpone cognitive decline. Traditional medical procedures such as magnetic resonance imaging are costly, involve long waiting lists, and require complex analysis. Alternatively, for the past years, researchers have successfully evaluated AD detection approaches based on machine learning and electroencephalography (EEG). Nonetheless, these approaches frequently rely upon manual processing or involve non-portable EEG hardware. These aspects are suboptimal regarding automated diagnosis, since they require additional personnel and hinder portability. In this work, we report the preliminary evaluation of a self-driven AD multi-class discrimination approach based on a commercial EEG acquisition system using sixteen channels. For this purpose, we recorded the EEG of three groups of participants: mild AD, mild cognitive impairment (MCI) non-AD, and controls, and we implemented a self-driven analysis pipeline to discriminate the three groups. First, we applied automated artifact rejection algorithms to the EEG recordings. Then, we extracted power, entropy, and complexity features from the preprocessed epochs. Finally, we evaluated a multi-class classification problem using a multi-layer perceptron through leave-one-subject-out cross-validation. The preliminary results that we obtained are comparable to the best in literature (0.88 F1-score), what suggests that AD can potentially be detected through a self-driven approach based on commercial EEG and machine learning. We believe this work and further research could contribute to opening the door for the detection of AD in a single consultation session, therefore reducing the costs associated to AD screening and potentially advancing medical treatment.

1. Introduction

The term dementia refers to a group of neurological pathologies dominated by a progressive loss of cognitive functions [1]. Among these pathologies, Alzheimer's disease (AD) represents between 60% and 70% of the cases, and affects more than fifty million people worldwide [2]. AD patients experience a decline in cognitive areas such as reasoning, memory, and orientation. Although this disease was firstly observed in 1906, its etiology remains unclear, and a definitive diagnosis can only be established upon brain autopsy. Nonetheless, researchers have identified two main hallmarks that start to form before the impairment is notable: amyloid plaques and neurofibrillary tangles [3,4]. Amyloid plaques are deposits of proteins that lose their standard structure and aggregate around the neurons. Alternatively, neurofibrillary tangles are thickened fibrils that encircle the nucleus of the neurons. In this context, mild cognitive impairment (MCI) is considered as a transitional stage between normal aging and AD. MCI patients experience minor memory losses, but they do not interfere with their daily-life activities. However, previous studies have found MCI patients progress to AD faster than healthy individuals of the same age [5]. Consequently, early detection is crucial to control the progression of the disease and defer the intellectual decline [6].

The purpose of AD detection techniques is to expose the cognitive and physical effects produced by the disease. Traditionally, this is accomplished through neuropsychological tests and medical procedures. Neuropsychological tests are evaluations designed to

* Corresponding author.
E-mail address: malg@ugr.es (M.Á. Lopez-Gordo).

https://doi.org/10.1016/j.cmpb.2022.106841

assess the intellectual areas affected in the AD course, like memory, orientation, and language. However, previous works have reported their lack of sensitivity and high variability [7,8]. Alternatively, medical procedures are designed to unfold the physical damages produced in specific brain areas. Among them, cerebrospinal fluid (CSF) analysis is the most reliable, since it allows the study of biomarkers linked to amyloid plaques and neurofibrillary tangles [9]. However, this fluid is obtained via lumbar puncture, a costly and invasive medical procedure not feasible for all patients. Conversely, medical imaging techniques enable the representation of static and functional images of the brain. These images are rendered using different procedures such as magnetic resonance imaging (MRI) [10], positron emission tomography (PET) [11], and single photon emission computed tomography (SPECT) [12]. Even though these techniques are accurate, they lack temporal resolution, involve long waiting lists, and their analysis is traditionally based on visual inspection. Therefore, these procedures represent an unsuitable alternative towards early detection.

Alternatively, for the past decade researchers have explored the use of electroencephalography (EEG) for AD detection [13–16]. EEG is a neurophysiology technique to acquire brain electrical signals through electrodes placed on the scalp. This technique is affordable, portable, and non-invasive. Hence, it represents a promising alternative for the detection of neurological diseases. In this regard, researchers have recently approached AD detection through the combination of EEG processing and machine learning algorithms. As the authors of [17] outline in their comprehensive review, these studies typically target two main goals: detection of AD cohorts and analysis of AD progression. Detection studies attempt to classify well defined clinical cohorts such as AD and MCI. Conversely, progression studies attempt to identify the participants who convert to AD after a longitudinal evaluation. In AD detection studies, such as the one presented in this paper, researchers typically extract features to characterize the three main AD effects on EEG activity: slowing, complexity reduction, and loss of synchronization [17–20]. These effects are quantified using multiple features, including: spectral and wavelet analysis (slowing) [21,22], entropy and information theory (complexity) [23–26], and coherence (synchronization) [15,27]. For this purpose, the usual methodological stages considered in state-of-the-art works include preprocessing, time-frequency and non-linear analysis, feature extraction, and classification. For instance, in [28], the authors investigated a feature selection system in the context of two binary classification tasks: MCI vs controls, and mild AD vs controls. Their results were promising as they obtained classification rates of 0.95 and 1, respectively, using synchrony and relative power features. Similarly, the authors of [21] applied cluster analysis using power and coherence measures to a similar classification task, yielding a classification accuracy of 0.91. More recently, in [29], the authors estimated statistical coefficients from a time-frequency map to perform two classification tasks: AD vs controls, and AD vs MCI, yielding accuracies of 0.96 and 0.87, respectively. In [30], the authors evaluated the classification of AD patients, frontotemporal dementia patients (FTD), and controls, from resting state EEG activity. The authors obtained 0.86 accuracy for the FTD vs controls task using a random forest classifier. Alternatively, they obtained 0.79 accuracy for the AD vs controls task using a decision tree classifier.

Although the study of AD progression is not the goal of the present work, we believe it is worth to note that multiple works have focused on this problem for the past years. Typically, the authors of these works perform an EEG acquisition follow-up from participants at risk of suffering AD. In [31], for instance, they performed an initial and a three-month follow-up evaluation of a group of MCI patients using high-density EEG. Then, the authors successfully applied a convolutional neural network (CNN) and power spectral density (PSD) features to classify the recordings into AD and MCI. They also analyzed which channels and frequencies resulted more active in the AD progression. Although deep learning algorithms, such as the CNN, have been successfully applied to AD cohort discrimination and progression analysis, these algorithms have not yet been applied on large EEG datasets [17].

The works referred before attempted to discriminate AD cohorts through binary classification. Conversely, a few authors have approached AD detection as a three-class classification problem. This kind of approach is essential for the simplification of the detection protocol. The first attempt was presented in [32], where the authors extracted complexity and spectral features to discriminate MCI, AD, and controls using a pairwise classification method. In [25], the authors trained a multilayer perceptron (MLP) to discriminate the three cohorts from spectral and non-linear features, yielding an accuracy of 0.63. In [33], the authors tackled a three-class classification task via deep learning. They proposed the application of a CNN to classify MCI patients, AD patients, and controls from 2D grayscale images obtained from the PSD of the EEG. Alternatively, although the most studied three-class classification task corresponds to MCI vs AD vs controls, the authors of [34] successfully analyzed the mild AD vs moderate AD vs controls task. More recently, in [35], the authors proposed a new Wavelet based analysis referred as lacsogram. They used cepstrum and lacsogram distance features to classify MCI patients, AD patients, and controls, and they obtained a classification accuracy of 0.96. Likewise, in [36], the authors successfully evaluated a similar classification task under different resting conditions.

In general, the works discussed in the previous paragraph approached three-class classification of AD cohorts with success. Nonetheless, they present certain features that hinder the self-driven detection of the disease. For instance, the need for visual inspection of the EEG signals, hardware-based artifact processing, or the use of clinical EEG acquisition systems that require trained personnel to be operated. In contrast, self-driven AD classification could contribute to reducing the costs associated to screening and help to advance medical treatment. In this context, the use of commercial EEG acquisition devices and the implementation of self-driven EEG processing pipelines may have the potential to speed up the analysis, foster reproducibility, and would not require additional personnel to be operated. Considering this, in this paper we report the preliminary results of a fully self-driven approach for AD three-class discrimination based on a commercial EEG acquisition device and self-driven processing. To evaluate our approach, we conducted a study alongside the Cognitive and Behavioral Neurology Unit (CBNU) at Hospital Universitario Virgen de las Nieves de Granada (Spain). We recorded the resting state brain activity of a group of MCI patients, AD patients, and controls, using a sixteen-channel wearable acquisition device. Then, we performed artifact rejection through automated independent component analysis (ICA) and Autoreject, a self-driven artifact rejection algorithm. We evaluated a multi-layer perceptron (MLP) for the three-class discrimination of the three study groups via leave-one-subject-out (LOSO) cross-validation. The preliminary results that we obtained are comparable to the best in literature. Although the sample size that we considered is reduced, our results suggest that AD and MCI could potentially be detected using a self-driven multiclass approach based on commercial wearable EEG. This is promising from the standpoint of screening protocols, since it could potentially enable the inclusion of AD detection techniques that are ubiquitous, affordable, and accurate. Nonetheless, further research must validate the conclusions yielded in this work.

2. Materials and methods

In this section, we review the main aspects of the study methodology. First, we describe the cohorts involved in this re-

Table 1
Group, sex, and age distribution of the participants involved in the study. The age column includes the mean age ± the standard deviation.

Group	Females	Males	Age
MCI-non-AD	0	6	72.8 ± 6.5
mild AD	7	4	68.3 ± 4.6
control	8	1	66.7 ± 3.4

search. Then, we disclose the experimental procedure, and we report the processing steps that we applied. Finally, we describe the feature extraction procedure and the implementation of the three-class classification.

2.1. Participants

The head of the CBNU recruited twenty-six participants for this study. Table 1 shows the participants group, sex, and age distributions. The participants labeled as mild AD and MCI-non-AD were CBNU patients at Hospital Universitario Virgen de las Nieves. These patients were diagnosed from PET- Amyloid or CSF analysis.

CSF analysis was carried out by two different laboratories during the recruitment period, and the reference cutoff value for the patients was that established by the corresponding laboratory. This cutoff value was based on a model of AD patients versus non-AD patients with no age stratification. CSF was obtained via lumbar puncture using a 20-gauge needle and syringe. The samples were collected in polypropylene collection tubes and sent immediately to the laboratory, where an ELISA Innotest assay was used to determine levels of Aβ42, total-τ, and τ-phosphorylated fraction. The results of the analysis were codified as normal or pathological.

PET-Amyloid was analyzed using 18F-florbetaben (FBB) by trained nuclear medicine specialists who had completed the learning curve for accredited PET-FBB scan interpretation and were blinded to the clinical situation of the patient. They reported the scans as positive (loss of grey-white matter contrast; regional cortical tracer uptake in any cortical target region: lateral temporal, frontal, posterior cingulate precuneus, or parietal), or negative (good grey-white matter contrast; no tracer uptake in target regions) for amyloid plaque presence. Cases with doubtful imaging results were discussed by both specialists to achieve a consensus.

These medical analysis were performed under the highest clinical guarantees, and we considered them as the gold standard for the patients diagnosis. Consequently, we labeled the patients as mild AD (CSF pathological results or PET positive amyloid plaque presence) and MCI-non-AD (CSF normal results or PET negative amyloid plaque presence). The rest of the participants were healthy age matched controls who did not report any neurological condition and were not CBNU patients. To conduct the study, we followed a protocol approved by the ethics committee at Hospital Universitario Virgen de las Nieves de Granada. Furthermore, the participants signed an informed consent before the start of the experiment, and clinical personnel supervised them throughout the entire session.

2.2. Setup and experimental procedure

The objective behind the engagement of the participants was the acquisition of their resting state EEG activity for our study. Nonetheless, taking advantage of the fact that the participants attended the consulting room for the acquisition, they performed a cognitive test consisting of two cognitive tasks. The details regarding such test are not reported in this work since they are out of the scope of this study, and they did not interfere with the EEG acquisition in any manner.

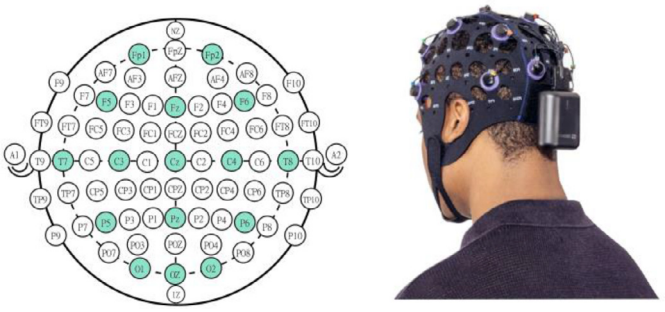

Fig. 1. (Left) Electrode setup selected for this study. We considered sixteen electrodes from the extended 10-20 International System to uniformly cover the scalp. (Right) Versatile semi-dry acquisition system used for the EEG data capture. The system operates at a fixed sampling rate of 256 Hz and provides support for sixteen electrodes.

Before the start of the experiment, we requested the participants to carefully read and sign the informed consent. Then, we acquired three minutes of their eye-open resting state EEG activity in three instants: prior to the first cognitive task, after this task, and after the second cognitive task. To prevent edge effects, we only studied the central two-minute segment of each recording. To perform our analysis, we concatenated the three recordings associated to each participant, hence, we considered a total EEG time of six minutes per participant.

To record the brain electrical activity, we used the Versatile wireless wearable system by Bitbrain. This commercial device includes a Bluetooth acquisition module working at 256 Hz, and an EEG headset with semi-dry electrodes. For the electrode setup, we considered sixteen sensors located at positions Fp1, Fp2, F5, Fz, F6, T7, T8, C3, Cz, C4, P5, Pz, P6, O1, Oz, and O2 of the extended 10-20 International System, and we referenced them to the left ear lobe. We selected this sensor montage to evenly cover the scalp following setups from analogous studies [21,37,38]. Fig. 1 represents the electrode setup selected in this study along with the Versatile system.

2.3. Signal processing

First, we filtered the raw EEG signals with a 1690 order bandpass FIR filter with 1-45 Hz passband and zero phase-shift. We selected a FIR filter rather than an IIR filter because our analysis did not entail high throughput restrictions, and we gave priority to filter control and stability [38]. Then, we split the filtered signals into four-second epochs without overlapping.

Subsequently, we implemented automated artifact rejection in two steps: Autoreject and ICA. First, we applied the Autoreject algorithm to automatically find an artifact threshold per channel and identify and reject bad data spans. Autoreject is an automated data-driven artifact rejection algorithm based on Bayesian optimization and cross-validation. The algorithm is available as a Python module and its application is straightforward. In terms of performance, Autoreject has been evaluated against several datasets and achieved equal or better performance compared to typical approaches. We refer the interested reader to [39], where the developers of the algorithm provide a comprehensive review of its implementation and evaluation. Subsequently, we applied ICA to correct blink artifacts. ICA enables the decomposition of a signal conformed by multiple sources into those sources and a mixing matrix. This is particularly useful to process EEG signals, as they represent the combination of internal brain sources. Based on this, ICA is typically used to detect and remove artifactual components such as blinks. These components display high variance and a spatial distribution towards the frontotemporal area of the scalp.

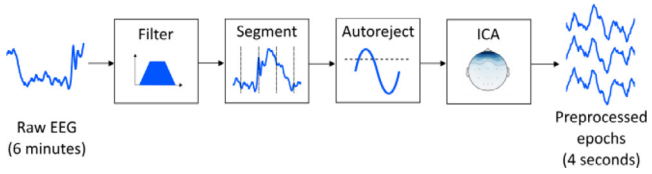

Fig. 2. Self-driven signal processing pipeline implemented in this study. Initially, we applied a 1-45 Hz bandpass FIR filter to remove the power line interference and retain the desired spectral content. Then, we segmented the filtered signals into four-second epochs without overlapping. Finally, we implemented automated artifact removal in two stages: Autoreject algorithm and ICA. The input and output signals represent the raw EEG and the preprocessed epochs of a single participant.

Therefore, to identify the blink artifact component, we used Fp1 as an electro-oculogram proxy, and automatically identified the component that showed the highest correlation with this sensor. Then, we removed the blink component from the source matrix, and we reconstructed the EEG signals. Fig. 2 shows the signal processing stage described in this subsection.

2.4. Feature extraction

After applying the EEG processing pipeline, we extracted three features per EEG channel from each preprocessed epoch: relative power (RP) in the five main EEG bands, spectral entropy (SE), and Hjorth complexity (HC). We examined these features because they have been already validated in analogous studies [21,32,40–42]. RP represents the fraction of the total signal power that is contained in a frequency range. Alternatively, SE represents the Shannon entropy of the power spectrum of the signal. SE describes the uniformity of the power spectrum distribution, and, hence, the irregularity of the EEG. Therefore, SE is minimal for a pure sine wave and maximal for white noise [43]. Finally, HC is one of the three Hjorth parameters (activity, mobility, and complexity), and is derived as the ratio of the mobility of the first derivative of the signal to the mobility of the signal itself [44]. The mathematical representation of the three features described in this paragraph are presented in Eqs. (1)-(3).

$$RP = \frac{\sum_{f_i}^{f_o} P}{\sum_{\forall f} P} \tag{1}$$

$$SE = -\sum_{f} S(f) * \log_2 S(f) \tag{2}$$

$$HC = \frac{\sigma_s'' / \sigma_s'}{\sigma_s' / \sigma_s} \tag{3}$$

In Eq. (1), f_i and f_o represent the lower and upper frequency bounds of a particular frequency band. For this study, we estimated RP in the five main EEG bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (> 30 Hz). In Eq. (2), f represents the frequencies in the frequency range of the signal, and S represents its normalized power spectrum. Finally, in Eq. (3), σ_s, σ_s', and σ_s'' represent the variance of the signal, the variance of its first derivative, and the variance of its second derivative.

After the feature extraction from the preprocessed epochs, we concatenated the features to create the feature matrix. Then, following the approach presented in [45], we averaged the features of every S adjacent rows in the feature matrix to increase the signal-to-noise ratio and reduce the size of the dataset. We examined four values for S: 6, 8, 10, and 12, and we kept the one that yielded the best performance.

2.5. Classification

Following feature extraction, we implemented a three-step classification pipeline including feature scaling, feature selection, and classification. First, the feature scaler normalizes the features between zero and one. This technique is extensively applied in machine learning to enhance the performance of non-tree-based models. Then, the feature selector gathers the most relevant features according to a predefined strategy. We suggest the interested reader to review [46] for a comprehensive review of frequent feature selection strategies. In this study, we applied a method based on the chi-squared test to identify the features most related to the target. We selected this strategy owing to its intrinsic speed. Lastly, for the last stage of the pipeline, we selected a MLP as this algorithm inherently allows multilabel classification.

To find the best combination of hyperparameters for the elements in the pipeline we applied grid-search. For the feature matrix averaging procedure described the last paragraph of Section 2.4, we evaluated different values for the number of rows to average. For the feature selection stage, we evaluated different values for the percentage of features selected. For the MLP, we evaluated different values for the architecture of the network, the L2 penalty (Alpha), and the activation function. Table 2 shows the different value ranges we considered during grid search for all the mentioned hyperparameters. We refer the interested reader to the scikit-learn documentation for a comprehensive description of the hyperparameters reported in this table. The best combination of hyperparameters is reported in the Results section.

Table 2
List of hyperparameters evaluated during grid search cross-validation[a].

Stage	Hyperparameter	Value range
Epoch averaging	Epochs averaged	6, 8, 10, 12
Feature selection	Feature percentage	10, 25, 50, 75, 100
Classification	Layer sizes	3-3-3, 4-4-4, 5-5-5
	Alpha	10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}
	Activation	relu, tanh

[a] Note that for the layer sizes parameter, each figure between the parenthesis refers to a hidden layer, and its value represents the number of neurons in the layer. For instance, 4-4-4 represents a MLP with three hidden layers and four neurons in each layer.

To evaluate the classification performance during the grid search, we applied cross-validation using a built-in function from scikit-learn Python module (GridSearchCV). This function evaluates the classification pipeline using different sets of hyperparameters and a cross-validation scheme. Since this study involves the discrimination of clinical groups, we followed a LOSO strategy. Thereby, data is split into as many folds as participants. Then, for each fold, the training set holds data from all the participants except one, whose data is reserved for the test set. Thus, information from a participant is never in the training set and the test set at the same time. This prevents from positive bias and represents a robust alternative to evaluate the general performance of the model. Fig. 3 displays the stages described in this section.

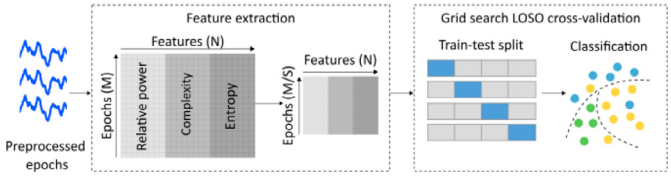

Fig. 3. Self-driven feature extraction and classification pipeline. First, we performed feature extraction on the preprocessed epochs. This procedure yielded a feature matrix with M rows (epochs) and N features. Then, we averaged every S consecutive epochs to improve the signal-to-noise ratio (we assessed S values of 6, 8, 10, and 12). Subsequently, we performed grid search cross-validation to obtain the best set of hyperparameters for the classification pipeline.

3. Results

In this section, we report the results we obtained for the three-class classification problem examined in this study. Owing to lack of cooperation (S09 and S18), presence of additional diseases (S12), and poor signal quality (S21 and S26), we excluded the data from five participants. Hence, we considered data from twenty-one participants in our analysis.

Fig. 4 represents the confusion matrix at the epoch level. To create this matrix, we gathered the true epoch labels and classifier predictions from the test sets used during cross-validation, and we estimated the distribution of the predictions across the three classes.

Fig. 5 shows the confusion matrix at the subject level. In this case, we evaluated the ability of the classifier to discriminate the participants via majority vote. This is, if the model correctly classified most of the epochs for a given participant, we considered it correctly discriminated the participant. Otherwise, we considered it misclassified the participant.

Table 3 presents a class report with the scores obtained by the model during cross-validation in terms of precision and recall. To create this table, we followed the procedure described for the epoch-level confusion matrix. We also included the average cross-validation F1-score.

Lastly, the best set of hyperparameters found via grid search is reported in Table 4.

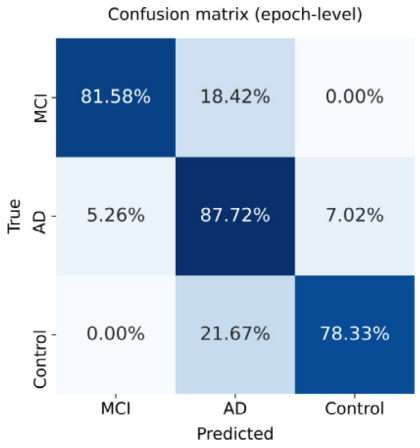

Fig. 4. Epoch-level confusion matrix for the discrimination of MCI-non-AD, mild AD, and controls.

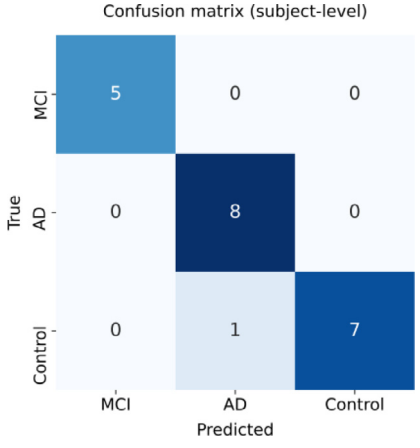

Fig. 5. Subject-level confusion matrix for the discrimination of MCI-non-AD, mild AD, and controls.

Table 3
Classification report for the mild AD vs MCI-non-AD vs control discrimination problema.

Cohort	Precision	Recall	F1-score
mild AD	0.91	0.82	0.88
MCI-non-AD	0.71	0.88	±
control	0.92	0.78	0.05

[a]We estimated the metrics displayed in this table after gathering all the test true labels and classifier predictions from the cross-validation. The right-most column represents the average cross-validation F1-score ± the standard error of the mean (SEM).

Table 4
Best set of hyperparameters found via grid search cross-validation.

Hyperparameter	Best value
Epochs averaged	12
Feature percentage	100
MLP Layer sizes	5-5-5
MLP Alpha	10^{-6}
MLP Activation	relu

4. Discussion

The purpose of this study was to evaluate the potential of a fully self-driven approach for AD three-class discrimination based on a commercial EEG acquisition device and automated processing. To this end, we recorded the resting state activity of a group of participants from these cohorts, and we developed a classification pipeline built on signal processing and machine learning. We used a commercial wearable EEG acquisition device that does not require trained experts to be operated. Thus, no further personnel aside from a neurologist is required to perform the data acquisition. The preliminary results that we obtained suggest that self-driven multiclass discrimination of AD cohorts from commercial EEG can be successfully performed. This may contribute to opening the door for the future implementation of affordable and portable early AD screening techniques. Nonetheless, further research is required to validate the preliminary results reported in this work.

In terms of epoch-level performance, the confusion matrix in Fig. 4 suggests that remarkable classification performance can be achieved following the approach presented in this work. Specifically, classification accuracy yielded by our classifier was 81.58%, 87.72%, and 78.33% for the MCI-non-AD, mild AD, and control epochs analyzed. Notably, the classifier yielded the best results for the mild AD and MCI-non-AD groups. This points at its ability to identify the patient cohorts, what is essential in the discrimination of clinical groups. This is further noted in the participant-level confusion matrix displayed in Fig. 5. According to this figure, our approach only misclassified a control participant. Naturally, the same conclusions are drawn from the cross-validation metrics reported in Table 3. As shown in this table, the classifier yielded an average F1-score of 0.88 ± 0.05.

According to the preliminary results obtained in this work, we believe a self-driven multi-class classification approach could present some advantages compared to the approaches reported in literature. For clarity, Table 5 summarizes the main aspects of related approaches. First, it is important to acknowledge that a direct comparison between these works is not entirely feasible due to the heterogeneity of the cohorts reported. Whilst most studies considered AD, MCI, and controls [16,32,36,47], other studies evaluated different clinical groups. For instance, [34] examined mild AD, AD, and controls, and [35] evaluated MCI, mild AD (ADM), advanced AD (ADA), and controls. In this sense, we believe the cohorts that we considered (MCI-non-AD, mild AD, and controls) are

Table 5
Comparison of studies on multi-class discrimination of AD cohorts. The columns indicate, from left to right, the authors, the cohorts involved, the diagnosis procedure to label the cohorts, the EEG acquisition device, the number of sensors utilized during the acquisition, the artifact rejection technique applied, and the performance metrics yielded for the multi-class discrimination of the cohorts.

Study	Cohorts	Diagnosis	EEG device	Sensors	Artifact rejection	Performance
McBride et al. [32]	17 early AD 16 MCI 15 HC	Cognitive tests and other evaluations	Neuroscan II	32	Not reported	F1-score = 0.83 (subject-level)
Ruiz-Gómez et al. [25]	37 AD 37 MCI 37 HC	NIA-AA	XLTEK Natus medical	19	Visual inspection	Accuracy = 0.63 (subject -level)
Tzimourtra et al. [34]	8 mild AD 6 AD 10 HC	MMSE	Nihon Kohden EEG 2100	19	Hardware	F1-score = 0.85 (epoch-level)
Ieracitano et al. [47]	63 AD, 63 MCI 63 HC	DSMMD	Not disclosed	19	Visual inspection	F1-score = 0.81 (epoch-level)
Oltu et al. [16]	8 AD 16 MCI 11 HC	MoCA	Nihon Kohden EEG 1200	19	Hardware	Accuracy = 0.94 (epoch-level)
Rodrigues et al. [35]	11 HC 8 MCI 11 ADM 8 ADA	MMSE	Not disclosed	19	Not specified	AUC = 0.95 (epoch-level)
Sharma et al. [36]	16 AD 16 MCI 15 HC	DSMMD	SOMNOscreen EEG 32	21	EEGLAB	F1-score = 0.85 (epoch-level)
Our approach	8 mild AD 5 MCI-non-AD 8 HC	CSF/PET	Versatile EEG 16	16	Autoreject and ICA	F1-score = 0.88 (epoch-level) Accuracy = 0.95 (subject -level)

the most suitable for early detection research. Another fundamental difference among these studies is the diagnosis procedure. Most previous works diagnosed the participants based on neuropsychological tests: [34,35] applied the mini mental state examination (MMSE), [16,36] used the Montreal cognitive assessment (MoCA), [36,47] applied the recommendations of the Diagnostic and Statistical Manual of Mental Disorders (DSMMD), [25] followed the guidelines of the National Institute of Aging and Alzheimer's Association (NIA-AA), and [32] applied a battery of cognitive tests and other evaluations. Conversely, we identified the participants based on the diagnosis yielded from a CSF analysis and a PET scan, medical procedures with higher reliability compared to neuropsychological examinations. In terms of electrode setup, most of these works selected 19-21 sensors [16,25,34–36,47], with only [32] using a considerably higher number of sensors (32). In this sense, we believe the use of a low number of electrodes promotes participant comfort and simplifies the montage. Regarding the acquisition system, we used the device with the highest degree of portability compared to the rest of approaches gathered in Table 5, as most of them utilized medical-oriented systems with reduced mobility.

Regarding artifact processing, [24] and [47] performed epoch manual selection based on visual inspection. This is unsuitable from the early detection standing point because it involves additional personnel and delays the analysis. Alternatively, [34] and [16] used the built-in capabilities of their acquisition hardware, the Nihon Kohden EEG system, to perform artifact rejection, what hinders the flexibility of their proposal. Instead, [36] used EEGLAB toolbox for MATLAB to perform artifact rejection. However, the authors do not specify if they used a manual or an automated procedure. Finally, [42] and [35] did not specify the artifact removal techniques applied. For this preliminary study, we applied Autoreject and automated ICA. We believe this is a more appropriate alternative for artifact rejection because the analysis remains self-driven, what is essential for early detection, and also fosters the reproducibility of the results yielded.

With respect to performance, a direct comparison is hindered by the heterogeneity of the cohorts and the metrics used to evaluate the results. The F1-score reported in these works ranges from 0.81 to 0.89 [32,34,36,47]. Other authors reported accuracy values between 0.63 and 0.94 [16,25,47]. Alternatively, [35] reported excellent results in terms of the area under the curve (AUC). As per the results reported in Table 3, a self-driven approach based on a commercial EEG device, such as the one introduced in this work, can achieve a performance that is comparable to the best results in literature.

5. Conclusions

In this study, we evaluated the potential of a fully self-driven approach for AD three-class discrimination using a commercial EEG acquisition system. For this purpose, we conducted a study involving participants from three cohorts: MCI-non-AD, mild AD, and healthy age-matched controls. First, we recorded their eye-open resting state brain activity, and then we implemented a self-driven pipeline including artifact rejection, feature extraction, and classification. For artifact rejection, we applied Autoreject, a data-driven algorithm, and ICA. Then, we extracted the relative power, Hjorth complexity, and spectral entropy from the clean epochs. Finally, we performed grid search cross-validation under a LOSO strategy.

In terms of performance, the preliminary results that we obtained are comparable to those reported in analogous studies, or even better in some cases. These results are promising since self-driven detection could potentially outperform existing approaches, promote early detection, contribute to reducing costs, and encourage the reproducibility of results. This may help reduce detection times associated with other traditional medical procedures like clinical EEG and nuclear medicine procedures. In this regard, early detection is paramount to control neural loss and defer cognitive decline. This is especially important in the case of middle-aged adults, since the impact of the disease in their social life and workplace is higher compared to older adults.

Alternatively, whilst the median number of participants in the studies discussed in Table 5 is forty-seven, we only considered twenty-one. Therefore, in future studies, we must evaluate the approach described in this paper on a larger sample to reinforce the conclusions drawn in this work. In addition, we are aware the participants involved in this and similar studies, do not strictly represent the typical patients who attend to neurology services. These patients, usually suffer from other pathologies apart from dementia. Hence, to effectively transfer EEG-based self-driven approaches into the clinical ecosystem, we must evaluate their performance on real-life participants. This way, in the future, patients may benefit from accurate, fast, and affordable detection approaches that help advance their medical treatments.

Declaration of Competing Interest

The authors of this work, whose names are listed in the manuscript, certify that they have no affiliations or involvement in any organization with any economic interest, or non-financial interest in the subject discussed in this manuscript.

Acknowledgments

This research was supported by project PGC2018-098813-B-C31 funded by the Spanish Ministry of Science, Innovation and Universities, by European Regional Development Funds, and by project B-TIC-352-UGR20, co-funded by the Operative Program FEDER 2014-2020 and the Economy, Universities and Science Office of the Andalusian Regional Government. The authors would like to especially thank the Cognitive and Behavioral Neurology Unit at Hospital Universitario Virgen de las Nieves de Granada and the participants involved in the study for their collaboration in this research. Funding for open access charge: Universidad de Granada/CBUA.

References

[1] M.J. Moore, C.W. Zhu, E.C. Clipp, Informal costs of dementia care Estimates from the national longitudinal caregiver study, J. Gerontol. B Psychol. Sci. Soc. Sci. 56 (4) (2001) S219–S228, doi:10.1093/geronb/56.4.S219.
[2] World Alzheimer Report, The State of the Art of Dementia Research: New Frontiers, New Front (2018) 48.
[3] D.P. Perl, Neuropathology of Alzheimer's Disease, Mount Sinai J. Med. 77 (1) (2010) 32–42, doi:10.1002/msj.20157.
[4] A. Serrano-Pozo, M.P. Frosch, E. Masliah, B.T. Hyman, Neuropathological alterations in Alzheimer Disease, Cold Spring Harb. Perspect. Med. 1 (1) (2011) a006189, doi:10.1101/cshperspect.a006189.
[5] R.A. Sperling, et al., Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, Alzheimer's Dementia 7 (3) (2011) 280–292, doi:10.1016/j.jalz.2011.03.003.
[6] M. Riemenschneider, N. Lautenschlager, S. Wagenpfeil, J. Diehl, A. Drzezga, A. Kurz, Cerebrospinal fluid tau and beta-amyloid 42 proteins identify Alzheimer disease in subjects with mild cognitive impairment, Arch. Neurol. 59 (11) (2002) 1729–1734, doi:10.1001/archneur.59.11.1729.
[7] G. Kuslansky, et al., Detecting dementia with the hopkins verbal learning test and the mini-mental state examination, Arch. Clin. Neuropsychol. (2004) 16.
[8] M.S. Mendiondo, J.W. Ashford, R.J. Kryscio, F.A. Schmitt, Modelling mini mental state examination changes in Alzheimer's disease, Stat. Med. 19 (11–12) (2000) 1607–1616 doi:10.1002/(SICI)1097-0258(20000615/30)19:11/12⟨1607::AID-SIM449⟩3.0.CO;2-O.
[9] R.J. Perrin, A.M. Fagan, D.M. Holtzman, Multimodal techniques for diagnosis and prognosis of Alzheimer's disease, Nature 461 (7266) (2009) 916–922, doi:10.1038/nature08538.
[10] S.H. Hojjati, A. Ebrahimzadeh, A. Khazaee, A. Babajani-Feremi, Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM, J. Neurosci. Methods 282 (2017) 69–80, doi:10.1016/j.jneumeth.2017.03.006.
[11] S.R. Meikle, F.J. Beekman, S.E. Rose, Complementary molecular imaging technologies: high resolution SPECT, PET and MRI, Drug Discov. Today 3 (2) (2006) 187–194, doi:10.1016/j.ddtec.2006.05.001.
[12] J.M. Górriz, F. Segovia, J. Ramírez, A. Lassl, D. Salas-Gonzalez, GMM based SPECT image classification for the diagnosis of Alzheimer's disease, Appl. Soft Comput. 11 (2) (2011) 2313–2325, doi:10.1016/j.asoc.2010.08.012.
[13] F.J. Fraga, G.Q. Mamani, E. Johns, G. Tavares, T.H. Falk, N.A. Phillips, Early diagnosis of mild cognitive impairment and Alzheimer's with event-related potentials and event-related desynchronization in N-back working memory tasks, Comput. Methods Programs Biomed. 164 (2018) 1–13, doi:10.1016/j.cmpb.2018.06.011.
[14] P.A.M. Kanda, E.F. Oliveira, F.J. Fraga, EEG epochs with less alpha rhythm improve discrimination of mild Alzheimer's, Comput. Methods Programs Biomed. 138 (2017) 13–22, doi:10.1016/j.cmpb.2016.09.023.
[15] C.T. Briels, D.N. Schoonhoven, C.J. Stam, H. de Waal, P. Scheltens, A.A. Gouw, Reproducibility of EEG functional connectivity in Alzheimer's disease, Alzheimer's Res. Therapy 12 (1) (2020) 68, doi:10.1186/s13195-020-00632-3.
[16] B. Oltu, M.F. Akşahin, S. Kibaroğlu, A novel electroencephalography based approach for Alzheimer's disease and mild cognitive impairment detection, Biomed. Signal Process. Control 63 (2021) 102223, doi:10.1016/j.bspc.2020.102223.
[17] K.D. Tzimourta, et al., Machine Learning Algorithms and Statistical Approaches for Alzheimer's Disease Analysis Based on Resting-State EEG Recordings: A Systematic Review, Int. J. Neural Syst. 31 (5) (2021) 2130002, doi:10.1142/S0129065721300023.
[18] C. Babiloni, et al., Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms, Int. J. Psychophysiol. 103 (2016) 88–102, doi:10.1016/j.ijpsycho.2015.02.008.
[19] A.H.H. Al-Nuaimi, E. Jammeh, L. Sun, E. Ifeachor, Complexity measures for quantifying changes in electroencephalogram in Alzheimer's disease, complex. 2018 (2018) 1–12, doi:10.1155/2018/8915079.
[20] N. Benz, et al., Slowing of EEG background activity in Parkinson's and Alzheimer's disease with early cognitive dysfunction,", Front. Aging Neurosci 6 (2014), doi:10.3389/fnagi.2014.00314.
[21] R. Wang, J. Wang, H. Yu, X. Wei, C. Yang, B. Deng, Power spectral density and coherence analysis of Alzheimer's EEG, Cogn Neurodyn 9 (3) (2015) 291–304, doi:10.1007/s11571-014-9325-x.
[22] V. Bairagi, EEG signal analysis for early diagnosis of Alzheimer disease using spectral and wavelet based features, Int. J. Inf. Tecnol. 10 (3) (2018) 403–412, doi:10.1007/s41870-018-0165-5.
[23] C. Jiang, Y. Li, Y. Tang, C. Guan, Enhancing EEG-based classification of depression patients using spatial information, in: IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29, 2021, pp. 566–575, doi:10.1109/TNSRE.2021.3059429.
[24] C. Coronel, et al., Quantitative EEG markers of entropy and auto mutual information in relation to MMSE scores of probable Alzheimer's disease patients, Entropy 19 (3) (2017) Art. no. 3, doi:10.3390/e19030130.
[25] S.J. Ruiz-Gómez, et al., Automated multiclass classification of spontaneous eeg activity in Alzheimer's Disease and mild cognitive impairment, Entropy 20 (1) (2018) 35, doi:10.3390/e20010035.
[26] M. Şeker, Y. Özbek, G. Yener, M.S. Özerdem, Complexity of EEG dynamics for early diagnosis of Alzheimer's disease using permutation entropy Neuromarker, Comput. Methods Programs Biomed. 206 (2021) 106116, doi:10.1016/j.cmpb.2021.106116.
[27] A.H. Meghdadi, et al., Resting state EEG biomarkers of cognitive decline associated with Alzheimer's disease and mild cognitive impairment, PLoS One 16 (2) (2021) e0244180, doi:10.1371/journal.pone.0244180.
[28] E. Gallego-Jutglà, J. Solé-Casals, F.-B. Vialatte, M. Elgendi, A. Cichocki, J. Dauwels, A hybrid feature selection approach for the early diagnosis of Alzheimer's disease, J. Neural Eng. 12 (1) (2015) 016018, doi:10.1088/1741-2560/12/1/016018.
[29] C. Ieracitano, N. Mammone, A. Bramanti, S. Marino, A. Hussain, F.C. Morabito, A time-frequency based machine learning system for brain states classification via EEG signal processing, in: 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1–8, doi:10.1109/IJCNN.2019.8852240.
[30] A. Miltiadous, et al., Alzheimer's disease and frontotemporal dementia: a robust classification method of EEG signals and a comparison of validation methods, Diagnostics 11 (8) (2021) Art. no. 8, doi:10.3390/diagnostics11081437.
[31] F.C. Morabito, C. Ieracitano, N. Mammone, An explainable Artificial Intelligence approach to study MCI to AD conversion via HD-EEG processing, Clin. EEG Neurosci. (2021) p. 15500594211063662, doi:10.1177/15500594211063662.
[32] J.C. McBride, et al., Spectral and complexity analysis of scalp EEG characteristics for mild cognitive impairment and early Alzheimer's disease, Comput. Methods Programs Biomed. 114 (2) (2014) 153–163, doi:10.1016/j.cmpb.2014.01.019.
[33] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, F.C. Morabito, A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings, Neurocomputing 323 (2019) 96–107, doi:10.1016/j.neucom.2018.09.071.
[34] K.D. Tzimourta, et al., EEG window length evaluation for the detection of Alzheimer's disease over different brain regions, Brain Sci. 9 (4) (2019) Art. no. 4, doi:10.3390/brainsci9040081.
[35] P.M. Rodrigues, B.C. Bispo, C. Garrett, D. Alves, J.P. Teixeira, D. Freitas, Lacsogram: a new EEG tool to diagnose Alzheimer's disease, IEEE J. Biomed. Health Inform. 25 (9) (2021) 3384–3395, doi:10.1109/JBHI.2021.3069789.
[36] N. Sharma, M.H. Kolekar, K. Jha, EEG based dementia diagnosis using multiclass support vector machine with motor speed cognitive test, Biomed. Signal Process. Control 63 (2021) 102102, doi:10.1016/j.bspc.2020.102102.
[37] N.N. Kulkarni, V.K. Bairagi, Extracting salient features for EEG-based diagnosis of Alzheimer's disease using support vector machine classifier, IETE J. Res. 63 (1) (2017) 11–22, doi:10.1080/03772063.2016.1241164.
[38] G. Fiscon, et al., Combining EEG signal processing with supervised methods for Alzheimer's patients classification, BMC Med. Inform. Decis. Mak. 18 (1) (2018) 35, doi:10.1186/s12911-018-0613-y.
[39] M. Jas, D.A. Engemann, Y. Bekhti, F. Raimondo, A. Gramfort, Autoreject: Automated artifact rejection for MEG and EEG data, Neuroimage 159 (2017) 417–429, doi:10.1016/j.neuroimage.2017.06.030.
[40] L.R. Trambaiolli, N. Spolaôr, A.C. Lorena, R. Anghinah, J.R. Sato, Feature selection before EEG classification supports the diagnosis of Alzheimer's disease, Clin. Neurophysiol. 128 (10) (2017) 2058–2067, doi:10.1016/j.clinph.2017.06.251.
[41] N. Houmani, et al., Diagnosis of Alzheimer's disease with Electroencephalography in a differential framework, PLoS One 13 (3) (2018), doi:10.1371/journal.pone.0193607.
[42] S.-S. Poil, W. de Haan, W.M. van der Flier, H.D. Mansvelder, P. Scheltens, K. Linkenkaer-Hansen, Integrative EEG biomarkers predict progression to Alzheimer's disease at the MCI stage, Front. Aging Neurosci. 5 (2013), doi:10.3389/fnagi.2013.00058.
[43] T. Inouye, et al., Quantification of EEG irregularity by use of the entropy of the power spectrum, Electroencephalogr. Clin. Neurophysiol. 79 (3) (1991) 204–210, doi:10.1016/0013-4694(91)90138-T.

[44] N. Päivinen, S. Lammi, A. Pitkänen, J. Nissinen, M. Penttonen, T. Grönfors, Epileptic seizure detection: a nonlinear viewpoint, Comput. Methods Programs Biomed. 79 (2) (2005) 151–159, doi:10.1016/j.cmpb.2005.04.006.

[45] F.J. Fraga, et al., Towards an EEG-based biomarker for Alzheimer's disease: Improving amplitude modulation analysis features, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 1207–1211, doi:10.1109/ICASSP.2013.6637842.

[46] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electric. Eng. 40 (1) (2014) 16–28, doi:10.1016/j.compeleceng.2013.11.024.

[47] C. Ieracitano, N. Mammone, A. Hussain, F.C. Morabito, A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia, Neural Netw. 123 (2020) 176–190, doi:10.1016/j.neunet.2019.12.006.