This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2021.3069789, IEEE Journal of Biomedical and Health Informatics

Lacsogram: a New EEG Tool to Diagnose Alzheimer's Disease

Pedro M. Rodrigues*, Bruno C. Bispo, Carolina Garrett, Dílio Alves, João P. Teixeira and Diamantino Freitas

***Abstract*—This work proposes the application of a new electroencephalogram (EEG) signal processing tool - the lacsogram - to characterize the Alzheimer's disease (AD) activity and to assist on its diagnosis at different stages: Mild Cognitive Impairment (MCI), Mild and Moderate AD (ADM) and Advanced AD (ADA). Statistical analyzes are performed to lacstral distances between conventional EEG subbands to find measures capable of discriminating AD in all stages and characterizing the AD activity in each electrode. Cepstral distances are used for comparison. Comparing all AD stages and Controls (C), the most important significances are the lacstral distances between subbands θ and α (p=0.0014$<$0.05). The topographic maps show significant differences in parietal, temporal and frontal regions as AD progresses. Machine learning models with a leave-one-out cross-validation process are applied to lacstral/cepstral distances to develop an automatic method for diagnosing AD. The following classification accuracies are obtained with an artificial neural network: 95.55% for All vs All, 98.06% for C vs MCI, 95.99% for C vs ADM, 93.85% for MCI vs ADM-ADA. In C vs MCI, C vs ADM and MCI vs ADM-ADA, the proposed method outperforms the state-of-art methods by 5%, 1%, and 2%, respectively. In All vs All, it outperforms the state-of-art EEG and non-EEG methods by 6% and 2%, respectively. These results indicate that the proposed method represents an improvement in diagnosing AD.**

***Index Terms*—Alzheimer's disease, Mild-cognitive impairment, diagnose, artificial neural networks, cepstrum, lacsogram.**

I. Introduction

In recent decades, especially in developed countries, the average human lifespan has progressively increased owing to advances in medical science and practice, improvements in public health and standards of living [1]. This extension of life expectancy has brought deep changes in the age structure of the world population, with a progressive decrease of young people and an increasing proportion of elderly people. The aging population has become a fact of scientific interest around the world for many reasons but one is because the elderly are most vulnerable to acquire certain degenerative diseases [2].

Alzheimer's disease (AD) is a chronic and neurodegenerative disorder that progressively affects all brain functions [1]. It is one of the most debilitating diseases of modern societies and the leading cause of dementia [1]. There are more than 50 million people with AD or other related diseases all over the world and, according to estimates of Alzheimer's disease International Organization, they will be around 132 million in 2050 [3]. The most common clinical features of this persistent brain disease are progressive memory loss, communication difficulties, cognitive impairment and decreased functional ability [4]. When the last two mentioned effects become severe, patients may be exposed to risk situations and the presence of caregivers is required. The type and level of care needed will change over time as the disease progresses.

The gradual progression of AD is usually categorized in four stages. The pre-dementia stage is known as Mild Cognitive Impairment (MCI) and is characterized by subtle symptoms where cognitive deficits are mainly limited to memory and patient's life activities are preserved [5]. MCI is considered a transitional stage between normal aging and AD. Subjects with MCI are at higher risk of developing AD, but only between 10% and 15% of the subjects affected by MCI actually develop AD [6]. The next two stages are Mild and Moderate AD (ADM). They are characterized by a significant increase of cognitive deficiencies and a marked loss of independence and, therefore, monitoring becomes necessary in these stages. The last stage is Advanced AD (ADA). During this stage, the caregiver's role will expand to full time as constant monitoring becomes strictly necessary because patients are unable to perform any kind of task [7].

Until today, AD has unknown cause. Several factors are often associated with a higher risk for developing AD, namely, family history, gender, heart diseases, stress, obesity, smoking, lower educational qualifications, head injuries and Down's syndrome, but aging is considered the main risk factor [4]. Moreover, it is known that senile plaques and neurofibrillary tangles in the medial temporal lobe and cortical areas are two pathological hallmarks of brains ravaged by AD [4]. In a healthy brain, protein fragments are broken down and eliminated. In AD, on the other hand, fragments accumulate to form insoluble plaques. Neurofibrillary tangles are formed especially by the Tau protein which is part of a microtubule structure. It happens that, in AD, the Tau protein is abnormal and causes the collapse of the microtubule structures. There-

Manuscript was accepted for publication in March 20th, 2021. This work was supported by National Funds from FCT - Fundação para a Ciência e a Tecnologia through projects UIDB/50016/2020 and UIDB/05757/2020 and by the Doctoral Program of Biomedical Engineering of Faculty of Engineering of the University of Porto. Asterisk (*) indicates the corresponding author.

*Pedro M. Rodrigues is with Catholic University of Portugal, CBQF - Centre for Biotechnology and Fine Chemistry – Associated Laboratory, Faculty of Biotechnology, Rua Diogo Botelho 1327, 4169-005 Porto, Portugal, e-mail: prodrigues@porto.ucp.pt.

Bruno C. Bispo is with the Department of Electrical and Electronic Engineering, Federal University of Santa Catarina, 88040-370, Florianópolis, SC, Brazil, e-mail: bruno.bispo@ufsc.br.

Carolina Garrett is with the Faculty of Medicine of the University of Porto and with the Neurological Unity of University Hospital Center of São João, Porto, Portugal, e-mail: garrett.mc51@gmail.com

Dílio Alves is with the Neurological Unity of University Hospital Center of São João, Porto, Portugal, e-mail: alvesdilio@gmail.com.

João P. Teixeira is with CEDRI and UNIAG of Polytechnic Institute of Braganca, Braganca, Portugal, e-mail: joaopt@ipb.pt.

Diamantino Freitas is with the Faculty of Engineering of the University of Porto, Porto, Portugal, e-mail: dfreitas@fe.up.pt.

fore, parts of the brain start to shrink because the brain's nerve cells die. In AD last stage, damage is generalized and the brain tissue volume decreases very significantly [8].

In addition to unknown cause, researchers have neither yet discovered a good treatment nor developed a reliable diagnosis method for AD [9]. Currently available therapies only soften and slow the symptoms progression [9]. The diagnosis is difficult and generally made by excluding other possible causes of dementia symptoms [9]. As AD leads to death, in average, eight years after being diagnosed [9], an early and accurate diagnosis method plays a key role in medical intervention to reduce brain damage, preserve daily functioning for longer and give the patient time to plan the future. And, as AD is the costliest chronic disease for society, the development of auxiliary diagnostic methods may lead to reducing costs associated with this illness. Therefore, AD represents a public health problem and a challenge for the scientific community to find solutions that minimize its socioeconomic effects [9].

In order to improve the early AD diagnosis, during the last years, several studies have been performed for detecting AD in the MCI and ADM stages. Studies carried out with biomarkers based on cerebrospinal fluid (CSF), peptide beta-amyloid 1-42 (Aβ_{1-42}), total tau protein (t-tau) and phosphorylated protein tau (p-tau) achieved promising results in discriminating healthy subjects and ADM patients. In [10], a sensitivity of 96.4% and a specificity of 76.9% were achieved using (Aβ_{1-42}). Combining (Aβ_{1-42}) and p-tau, a sensitivity of 94% was obtained in [11]. By means of t-tau and p-tau at threonine 231 (p-tau$_{231P}$) or at threonine 181 (p-tau$_{181P}$), diagnostic sensitivity and specificity levels between 80% and 90% were obtained in [12]. The use of CSF markers allowed to achieved a sensitivity of 80% and a specificity of 90% in discriminating between ADM patients and control subjects as well as a sensitivity of 85% and a specificity of 75% in distinguishing between ADM patients, MCI patients and controls [13].

Studies performed with biomarkers based on neuroimaging techniques have also presented good results in discriminating between healthy subjects and ADM patients. In [14], magnetic resonance imaging (MRI) images were used to evaluate differences in hippocampal volume, achieving sensitivity, specificity and accuracy values equal to 84%. By means of positron emission tomography (PET) scans using Florbetaben (18F) amyloid-β, a sensitivity of 80% and a specificity of 91% were obtained in [15]. The association of PET scans with the biomarker [(11) C] Pittshburgh Compound B (PIB) proved to be a good tool to support AD diagnosis as it achieved a sensitivity of 100% and a specificity of 85% in discriminating healthy subjects and ADM patients [16], and a sensitivity of 56% and a specificity of 93% in distinguishing between controls and MCI patients [13]. Similarly, the use of 8F-2-fluoro-2-deoxy-d-glucose (FDG) PET images obtained sensitivity and specificity equal to 93% in distinction between controls and ADM patients [17], and an accuracy of 94% in discriminating between ADM patients, MCI patients and controls [13].

Despite the aforementioned studies have shown good diagnostic accuracy, nowadays the use of biomarkers for routine AD diagnostic purposes is not recommended in any consensus guidelines [18]. So far, clinical diagnosis of AD is typically based on the criteria established by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's disease and Related Disorders Association (NINCDS-ADRDA) [19]. They are based on the study of medical history, clinical examination, neuropsychological testing and laboratory assessments. The most important criterion established by NINCDS-ADRDA is the Mini Mental State Examination (MMSE), which allows to estimate the severity of cognitive loss. Accumulating evidence has shown that this criterion has been useful, achieving a sensitivity of 81% and specificity of 70% [20]. Furthermore, the electroencephalogram (EEG) is recommended by the EFNS guidelines 2010 as a tool to complement the clinical diagnosis [21].

The EEG has been used for several decades as a diagnostic tool for dementia. It is a physiological technique that records the spontaneous electrical brain activity originated by the neurons with high resolution through electrodes connected to the scalp [22]. As AD affects the nerve cells activity, the EEG may be helpful in AD identification. EEG recording systems are nowadays widely extended in clinical settings because they are inexpensive, non-invasive, mobile and fast. The EEG signal is typically divided into several frequency bands, such as delta (1-4 Hz, δ), theta (4-8 Hz, θ), alpha (8-13 Hz, α), beta (13-30 Hz, β) and gamma (30-40 Hz, γ) [22].

AD seems to affect EEG power in these subbands. Many studies have shown that AD causes a "slow down" on EEG signals, increasing the power at low frequencies (δ and θ) and decreasing the power at high frequencies (α and β) [8], [23]. This phenomenon, also called "shift-to-the-left", is characterized by the shifting of the power spectrum peak at the subband α to the lower subbands and is more notable in intermediate and advanced stages of AD [8], [23]. These changes are related to the destruction of cholinergic synapses in the Meynert nucleus, where the transferase responsible for the synthesis of acetylcholine is produced. A lack of acetylcholine is, in turn, responsible for disorders in the synchronization of synaptic potentials and, thus, pathological activity reveals itself in the form of slower EEG waves [8], [23].

Nowadays, there are a quite few EEG-based approaches with machine learning algorithms for differentiating AD stages. The features extracted from the EEG signals, the applied algorithms for feature selection and classification, and the accuracies of the state-of-art EEG-based methods are presented in the Table I. The use of cross-validation processes in the machine learning model evaluation is also indicated.

It is observed that all works performed binary classifications, usually between a AD stage and control group (C). The maximum accuracy for the pairs C vs MCI, C vs ADM and MCI vs ADM-ADA are 93%, 95% and 92%, respectively. On the other hand, only three studies performed a multi-class classification including all stages of AD and controls subjects (All vs All). The maximum accuracy for this case is 90%.

This work proposes a new method based on the wavelet transform with logarithmic transform, named lacsogram, that uses distances in this new domain to characterize and diagnose AD in the MCI, ADM and ADA stages through EEG signals. The characterization is made through a sensor level topographic map that traces the regional abnormalities caused by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2021.3069789, IEEE Journal of Biomedical and Health Informatics

Table I
SUMMARY OF THE STATE-OF-ART EEG-BASED METHODS FOR AUTOMATIC CLASSIFICATION OF AD.

Authors	Extracted Features	Classifier	Feature selection	Cross-validation	Accuracy (%)
Akrofi *et al.* [24]	Power and coherence features	Gaussian mixture model	Sequential forward floating search	Yes	C vs MCI: 90% All vs All: 77%
Afshari *et al.* [25]	Graph theory features	Linear discriminant	No	Yes	C vs ADM: 94%
Aghajani *et al.* [26]	Logarithmic power in subbands α, β, δ and θ	SVM	Singular value decomposition	Yes	C vs MCI: 84%
Besthorn *et al.* [27]	Coherence features	Linear discriminant	No	Yes	C vs ADM: 86%
Buscema *et al.* [28]	Ifast model features	Associative ANN	No	No	MCI vs ADM-ADA: 92%
Cassani *et al.* [29]	Subbands power and amplitude modulation rate-of-change	SVM	ANOVA test	Yes	C vs ADM: 79%
Huang *et al.* [30]	Global field power of subbands	Linear discriminant	ANOVA test	Yes	C vs ADM: 84% MCI vs ADM-ADA: 78% All vs All: 78%
Khatun *et al.* [31]	Subbands power	SVM	Candidate feature vector selection	Yes	C vs MCI 88%
Knott *et al.* [32]	Subbands power	Linear discriminant	ANOVA test	Yes	C vs ADM: 75%
Melissant *et al.* [33]	Independent component analyses features and subband power features	Bayes classifier, ANN, linear discriminant and k-nearest neighbour	Dimension reduction per channels	Yes	C vs AD: 94%
Petrosian *et al.* [34]	Subband signals	Recurrent ANN	ANOVA test	No	C vs ADM: 90%
Poil *et al.* [35]	Non-linear parameters (e.g. kurtosis, skewness, detrended fluctuation index)	Logistic regression	Genetic algorithm	No	MCI vs ADM-ADA: 85%
Viallate *et al.* [36]	Bump models	ANN	Principal component analysis	Yes	C vs MCI 93%
Rodrigues *et al.* [37]	Subbands power, spectral ratios and number of minima, maxima and zero crossing	ANN	Kruskal-Wallis test	Yes	C vs MCI: 77%; C vs ADM: 95%; MCI vs ADM-ADA: 83%; All vs All: 90%

AD in the different stages. The discriminative capacity of the proposed method is compared with a cepstrum-based method.

This paper is organized as follows: in section II, the EEG database is described; the signal processing methodology, including the proposed multiband lacstral analysis and the associated distance measures, is presented in section III; in section IV, statistical analyses and classification procedures are performed in order to develop a method for diagnosing AD in all stages, and the obtained results are presented and discussed; finally, the conclusions are presented in section V.

II. EEG DATASET

The dataset contains EEG recordings from 38 subjects: 11 healthy subjects as Controls (C), 8 patients with MCI, 11 patients with ADM and 8 patients with ADA. The average age and average MMSE score of the participants in each study group are presented in Table II. All participants gave their consent prior to participating in this study, whose protocol number CES198-14 was approved and authorized on 20 March 2015 by the local Ethics Committee of the University Hospital Center of São João (UHCSJ), Porto, Portugal.

The EEG samples were recorded from the 19 scalp loci of the International 10-20 configuration using a digital electroencephalograph in UHCSJ, Porto, Portugal. The participants were in a state of relaxation and with eyes closed. The sampling frequency was 256 Hz. The recordings were filtered through a digital band-pass filter with cut-off frequencies of 1 and 40 Hz, thus passing the frequency range of the conventional EEG subbands. The EEG channel signals were split into non-overlapped artifacts-free 5 s-long segments (1280 samples), as done in [37], [38], [39].

Table II
INFORMATION ABOUT THE EEG DATASET.

	Control Subjects	MCI Patients	ADM Patients	ADA Patients
#	11	8	11	8
Age average	74	80	79	79
MMSE average	28.68	26.29	18.89	11.50

III. SIGNAL PROCESSING METHODOLOGY

The proposed signal processing methodology is composed of 3 steps: EEG signal multiband processing, features extraction and features normalization. The signal processing methodology is depicted in Figure 1.

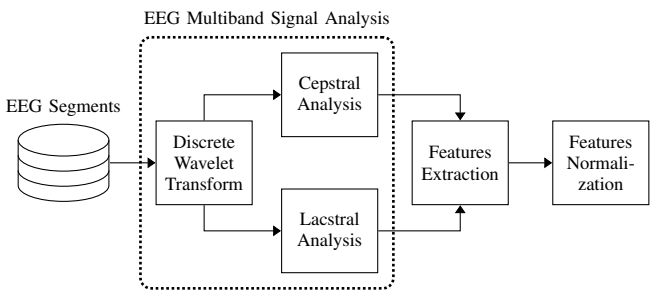


Figure 1. Signal processing methodology.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2021.3069789, IEEE Journal of Biomedical and Health Informatics

A. Multiband Decomposition with Discrete Wavelet Transform

The discrete wavelet transform (DWT) of a discrete-time finite-energy signal is its decomposition in a set of basis functions obtained from a finite number of prototype sequences and their time shifted versions [40]. It is an optimal tool for time-frequency signal analysis [40].

This structured expansion and its corresponding reconstruction are implemented by means of an octave-band critically decimated filter bank [40], [41]. Considering only the positive frequencies, the mth subband is confined to

$$W_k = \begin{cases} \left[0, \pi/2^S\right], & m = 0 \\ \left[\pi/2^{S-m+1}, \pi/2^{S-m}\right], & m = 1, 2, \ldots, S. \end{cases} \quad (1)$$

where $S+1$ is the number of subbands and π is the normalized angular frequency.

The DWT uses a synthesis scale function $\phi_1(n)$ and a synthesis wavelet function $\psi_1(n)$ defined as [41]

$$\phi_1(n) = g_{\mathrm{LP}}(n) \quad (2)$$

and

$$\psi_1(n) = g_{\mathrm{HP}}(n), \quad (3)$$

where $g_{\mathrm{LP}}(n)$ and $g_{\mathrm{HP}}(n)$ are the impulse responses of the half-band low-pass and high-pass synthesis filters, respectively.

Defining the following recursion formulas [41]

$$\phi_{i+1}(n) = \phi_i(n/2) * \phi_1(n) \quad (4)$$

$$\psi_{i+1}(n) = \psi_i(n/2) * \phi_1(n), \quad (5)$$

where the symbol $*$ denotes the convolution operator, the equivalent synthesis filter of the mth subband is given by [41]

$$g_m(n) = \begin{cases} \phi_S(n), & m = 0 \\ \psi_{S+1-m}(n), & m = 1, 2, \ldots, S. \end{cases} \quad (6)$$

The synthesis subband filters $g_1(n)$ to $g_S(n)$ are roughly dilated versions of the wavelet $\psi_1(n)$, thereby having approximately constant shapes, and $g_0(n)$ is a roughly dilated version of the scaling function $\phi_1(n)$, which is the basic low-pass filter of each reconstruction stage [41].

The DWT uses an analysis scale function $\tilde{\phi}_1(n)$ and an analysis wavelet function $\tilde{\psi}_1(n)$ defined as

$$\tilde{\phi}_1(n) = h_{\mathrm{LP}}(n) \quad (7)$$

and

$$\tilde{\psi}_1(n) = h_{\mathrm{HP}}(n), \quad (8)$$

where $h_{\mathrm{LP}}(n)$ and $h_{\mathrm{HP}}(n)$ are the impulse responses of the half-band low-pass and high-pass analysis filters, respectively.

Defining the following recursion formulas

$$\tilde{\phi}_{i+1}(n) = \tilde{\phi}_i(n/2) * \tilde{\phi}_1(n) \quad (9)$$

$$\tilde{\psi}_{i+1}(n) = \tilde{\phi}_i(n) * \tilde{\psi}_1(n/2^i), \quad (10)$$

the equivalent analysis filter of the mth subband is given by

$$h_m(n) = \begin{cases} \tilde{\phi}_S(n), & m = 0 \\ \tilde{\psi}_{S+1-m}(n), & m = 1, 2, \ldots, S. \end{cases} \quad (11)$$

If the wavelet base is biorthogonal, then the filter bank is perfect reconstruction [40], [41]. If the wavelet base is orthogonal, then the filter bank is of a perfect reconstruction type and the analysis filters are identical to the time-reversed synthesis filters [40], [41], that is,

$$h_m(n) = g_m(L_m - 1 - n), \quad (12)$$

where L_m is the length of the mth subband filter.

The mth subband signal, $m = 0, 1, \ldots, S$, is given by

$$x_m(n) = \begin{cases} \sum_{k=-\infty}^{\infty} x(k) h_m(2^S n - k), & m = 0 \\ \sum_{k=-\infty}^{\infty} x(k) h_m(2^{S-m+1} n - k), & m = 1, 2, \ldots, S. \end{cases} \quad (13)$$

In this work, the DWT is applied to each EEG segment in order to decompose them into the conventional EEG subbands, i.e., δ (1-4 Hz), θ (4-8 Hz), α (8-13 Hz), β (13-30 Hz) and γ (30-40 Hz). In this case, $S = 5$ and the subband signals are related to the EEG subbands as follows: $x_\delta(n) = x_0(n)$, $x_\theta(n) = x_1(n)$, $x_\alpha(n) = x_2(n)$, $x_\beta(n) = x_3(n)$ and $x_\gamma(n) = x_4(n)$. The signal $x_5(n)$ is not used. The Biorthogonal 3.5 is used as mother wavelet since it has proven to fit well with EEG signals from AD patients in [37], [38], [39]. The subband signals are upsampled to the original sampling frequency through the wavelet-based interpolation method [42].

B. Multiband Cepstral Analysis

Cepstral analysis is a nonlinear signal processing technique based on a homomorphic transformation that maps convolution in addition, resulting in cepstrum. It was proposed in 1963 by Bogert, Healy and Tukey as an alternative to the autocorrelation function for detecting echoes in seismic signals [43]. The cepstrum is useful to separate source and filter components.

The real cepstrum of a discrete-time finite-energy signal $x(n)$ is defined as [43]

$$x^c(n) = \mathcal{F}^{-1}\left\{\log\left[\left|\mathcal{F}\left\{x(n)\right\}\right|\right]\right\}, \quad (14)$$

where $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ denote the discrete-time Fourier transform (DTFT) and its inverse, respectively. In practice, an N-point fast Fourier Transform (FFT) is used in place of the DTFT. Although any base can be used in this computation, the decimal logarithm was employed in this work.

In the multiband cepstral analysis proposed in the present work, the real cepstrum is calculated for each subband signal $x_i(n)$, $i = \{\delta, \theta, \alpha, \beta, \gamma\}$, of each EEG segment using an 1280-point FFT, resulting in the cepstra $x_i^c(n)$, $i = \{\delta, \theta, \alpha, \beta, \gamma\}$, for each segment.

C. Multiband Lacstral Analysis

The scalogram is the time-frequency representation of the squared magnitude of the wavelet transform [44]. The proposed lacstral analysis brings the cepstrum computation closer to the scalogram. The lacsogram of a discrete-time finite-energy signal $x(n)$ is then defined as the squared magnitude of the inverse DWT of the signal's DWT coefficients logarithm magnitude, i.e.,

$$x^l(n) = \left|\sum_{m=0}^{S} y_m(n)\right|^2, \quad (15)$$

where

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2021.3069789, IEEE Journal of Biomedical and Health Informatics

$$y_m(n) = \begin{cases} \sum_{k=-\infty}^{\infty} |\log[x_m(k)]|\, g_m(n-2^S k), & m=0, \\ \sum_{k=-\infty}^{\infty} |\log[x_m(k)]|\, g_m(n-2^{S-m+1}k), & m=1,\ldots,S. \end{cases} \quad (16)$$

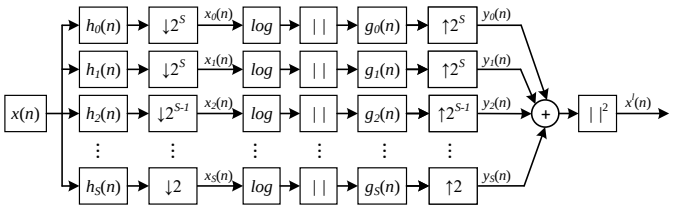


Figure 2. Illustration of the lacsogram computation.

The lacsogram computation is illustrated in Figure 2.

In the multiband lacstral analysis proposed in the present work, the lacsogram is calculated for each subband signal $x_i(n), i = \{\delta, \theta, \alpha, \beta, \gamma\}$, of each EEG segment using $S = 1$, resulting in the lacsograms $x_i^l(n), i = \{\delta, \theta, \alpha, \beta, \gamma\}$, for each segment. Although the lacstral analysis can be applied to any number of subbands, $S = 1$ is chosen for its first application.

D. Cepstral/Lacstral Distance Measures

Several measures can be calculated from the cepstrum and lacsogram. In the present work, cepstral/lacstral distances are calculated for the 5s-long segments of the 19 EEG channels of each subject. They are described below considering $i = \{\delta, \theta, \alpha, \beta, \gamma\}$ and $j = \{\delta, \theta, \alpha, \beta, \gamma\}$ as the subband indexes ($i \neq j$), $t = \{c, l\}$ as the multiband analysis type and N as the cepstrum/lacsogram signal length.

- The normalized root mean square (RMS) distance is defined as [45]

$$D1_{i,j}^t = l \times \sqrt{\left[x_i^t(1) - x_j^t(1)\right]^2 + p \times \sum_{n=2}^{N} \left[x_i^t(n) - x_j^t(n)\right]^2}, \quad (17)$$

  where $l = 4.3429$ and $p = 2$ are normalization factors [45]. The weight of the first cepstral/lacstral component is reduced because, in general, it does not contain any important information about the EEG signal [45].
- The normalized RMS distance without the first component is defined as [39]

$$D2_{i,j}^t = l \times \sqrt{p \times \sum_{n=2}^{N} \left[x_i^t(n) - x_j^t(n)\right]^2}. \quad (18)$$

  This distance also removes the first cepstral/lacstral component because, as mentioned, it usually does not contain any important information about the EEG signal.
- The Euclidean cepstral/lacstral distance is defined as [46]

$$D3_{i,j}^t = \sqrt{\sum_{n=1}^{N} \left[x_i^t(n) - x_j^t(n)\right]^2}. \quad (19)$$

  As the influence of the first cepstral/lacstral component is small, $D3_{i,j}^l$ is practically proportional to $D2_{i,j}^l$.
- The quefrency weighted cepstral/lacstral distance is defined as [47]

$$D4_{i,j}^t = \sqrt{\sum_{n=1}^{N} n \times \left[x_i^t(n) - x_j^t(n)\right]^2}. \quad (20)$$

- The quefrency rooted weighted cepstral/lacstral distance is defined as [46]

$$D5_{i,j}^t = \sqrt{\sum_{n=1}^{N} \sqrt{n} \times \left[x_i^t(n) - x_j^t(n)\right]^2}. \quad (21)$$

- The quefrency squared weighted cepstral/lacstral distance is defined as [46]

$$D6_{i,j}^t = \sqrt{\sum_{n=1}^{N} n^2 \times \left[x_i^t(n) - x_j^t(n)\right]^2}. \quad (22)$$

The weighting factors in $D4_{i,j}^t$, $D5_{i,j}^t$ and $D6_{i,j}^t$ aims to improve the prediction accuracy of the distances and, consequently, to find more abrupt differences between study groups comparing with the Euclidean distance $D3_{i,j}^t$ [46].

E. Normalization of the Cepstral/Lacstral Distance Measures

For each subband pair $\{i, j\}$, time/segment series of the distance measures $Dk_{i,j}^t$, $k = \{1, 2, 3, 4, 5, 6\}$, are determined for all EEG channels of each subject. In order to optimize execution time and performance of the classification procedures described later, the distance measures are normalized.

For each set $\{i, j, t, k\}$, the values of the distance measures are normalized through the Z-score normalization as [48]

$$Dk_{i,j}^t = \frac{Dk_{i,j}^t - \mu}{\sigma}, \quad (23)$$

where μ and σ denote mean and standard deviation of $Dk_{i,j}^t$, respectively. It is noteworthy that the normalization includes all segments of all EEG channel signals of all subjects.

IV. Statistical Analysis and Classification

In this work, the normalized values of the distance measures are used to perform statistical analyses and classification procedures. The statistical analysis include a global and an individual channel analysis. The global analysis aims to find distance measures with greater capacity of discriminating the study groups C, MCI, ADM and ADA. The channel analysis uses the distance measures selected from the global analysis in order to find evidence of activity and progression of AD in each electrode/channel, that is, throughout the scalp. The classification procedures do combine some distance measures and machine learning techniques to develop a method for diagnosing AD in its different stages.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2021.3069789, IEEE Journal of Biomedical and Health Informatics

A. Statistical Analysis

In the global analysis, for each set $\{i, j, t, k\}$, the normalized values of $Dk^t_{i,j}$ are averaged and grouped to be used as features. First, a double average procedure is applied. The normalized values are averaged across time/segments and channels, resulting in one mean normalized value per subject. Then, the mean normalized values are separated by AD stages, that is, C, MCI, ADM and ADA, and treated as distributions.

The normality and homoscedasticity of the distributions are assessed with the Kolmogorov-Smirnov and Levene tests, respectively. The hypothesis of parametric tests is not met. Thereafter, the Kruskal-Wallis test is used to determine if the null hypothesis that the data of the study groups (C, MCI, ADM and ADA) come from the same distribution is accepted. In this test, p-values lower than 0.05 indicate that there is a significant difference between the distributions and then the null hypothesis is rejected [49]. It is worthy mentioning that p-values are corrected for this multiclass comparison by the Bonferroni method [50]. The p-values corresponding to significant differences between the study groups are shown in the Table III, where the lowest value for each distance measure is highlighted in bold and n.s. means not significant.

Table III
$p-$VALUES OF THE DISTANCE MEASURES $Dk^t_{i,j}$ FOR BOTH MULTIBAND ANALYSIS METHODS.

Subband Indexes	Multiband Analysis	Features					
		$D1^t_{i,j}$	$D2^t_{i,j}$	$D3^t_{i,j}$	$D4^t_{i,j}$	$D5^t_{i,j}$	$D6^t_{i,j}$
i=δ j=θ	**Cepstral** $(t=c)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	Lacstral $(t=l)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
i=δ j=α	**Cepstral** $(t=c)$	0.0102	0.006	0.006	n.s.	n.s.	**0.0082**
	Lacstral $(t=l)$	0.0034	0.0034	0.0034	0.0032	0.0032	n.s.
i=δ j=β	**Cepstral** $(t=c)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	Lacstral $(t=l)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
i=δ j=γ	**Cepstral** $(t=c)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	Lacstral $(t=l)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
i=θ j=α	**Cepstral** $(t=c)$	0.0312	0.0166	0.0166	n.s.	n.s.	0.0096
	Lacstral $(t=l)$	**0.0014**	**0.0014**	**0.0014**	**0.0016**	**0.0020**	n.s.
i=θ j=β	**Cepstral** $(t=c)$	0.03	0.0456	0.0456	n.s.	n.s.	n.s.
	Lacstral $(t=l)$	0.0032	0.003	0.003	0.0042	0.0036	n.s.
i=θ j=γ	**Cepstral** $(t=c)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	Lacstral $(t=l)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
i=α j=β	**Cepstral** $(t=c)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	Lacstral $(t=l)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
i=α j=γ	**Cepstral** $(t=c)$	n.s.	n.s.	n.s.	n.s.	n.s.	0.0456
	Lacstral $(t=l)$	n.s.	n.s.	n.s.	n.s.	n.s.	0.0354
i=β j=γ	**Cepstral** $(t=c)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	Lacstral $(t=l)$	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

From Table III, it can be observed that only 28 of the 120 analyzed features, equivalent to 23.33%, are significant (p-value lower than 0.05). The significant features are, in large majority, the distance measures calculated with the subband pairs $\{\delta, \alpha\}$, $\{\theta, \alpha\}$ and $\{\theta, \beta\}$. This result shows the EEG waveform slow down, where the EEG power moves towards the low frequencies as AD progresses. A possible hypothesis to explain this phenomenon involves a loss of neurotransmitter acetylcholine. Agnoli *et al.* [51] reported that, after the administration of cholinergic drugs, the memory was improved and the EEG signals exhibited a shift tendency to normal values. The opposite effect occurs after the administration of anticholinergic drugs. Another study carried out by Knott *et al.* [32] reported a power decrease at subbands δ and θ as well as a power increase at subbands α and β, after administration of acetylcholine agonist nicotine to AD patients. The cognitive dysfunction of AD patients showed that there is a correlation between the EEG slow down and the appearance of AD [8].

It can also be noticed that $D1^l_{\theta,\alpha}$, $D2^l_{\theta,\alpha}$ and $D3^l_{\theta,\alpha}$, provided by the multiband lacstral analysis, present the best discriminant capacity, achieving a p-value of 0.0014. For the same subband pair, the distance measures $D2$ and $D3$ present the same p-value. This result demonstrates that the influence of the first cepstral/lacstral component, which is not included in $D3$, in the cepstral/lacstral distance measures is actually very small for EEG signals as discussed in Section III-D.

Compared with the multiband cepstral analysis, the multiband lacstral analysis makes 2 features no longer significant, namely, $D6^l_{\delta,\alpha}$ and $D6^l_{\theta,\alpha}$. In fact, for the distance $D6$, the cepstral analysis present the best discriminant results in $D6^c_{\delta,\alpha}$ and $D6^c_{\theta,\alpha}$. On the other hand, the lacstral analysis makes 6 features become significant, namely, $D4^l_{\delta,\alpha}$, $D4^l_{\theta,\alpha}$, $D4^l_{\theta,\beta}$, $D5^l_{\delta,\alpha}$, $D5^l_{\theta,\alpha}$ and $D5^l_{\theta,\beta}$. Thus, the lacstral analysis provides a greater amount of significant features, 16-12 over the cepstral analysis, and, most important, the 16 lowest p-values and the lowest p-value for 5 of the 6 distance measures, namely, $D1$, $D2$, $D3$, $D4$ and $D5$. This is quite relevant.

Therefore, it can be concluded that the proposed multiband lacstral analysis achieves better discriminant results in the vast majority of cases. The lowest p-value for each distance measure is obtained in $D1^l_{\theta,\alpha}$, $D2^l_{\theta,\alpha}$, $D3^l_{\theta,\alpha}$, $D4^l_{\theta,\alpha}$ and $D5^l_{\theta,\alpha}$ through the multiband lacstral analysis and in $D6^c_{\delta,\alpha}$ through the multiband cepstral analysis. This result indicates not only the advantage of the proposed lacstral analysis but also the importance of the subbands θ and α to discriminate AD stages.

Boxplots of the distributions of $D1^l_{\theta,\alpha}$, $D2^l_{\theta,\alpha}$, $D3^l_{\theta,\alpha}$, $D4^l_{\theta,\alpha}$, $D5^l_{\theta,\alpha}$ and $D6^c_{\delta,\alpha}$ are shown in Figure 3. Except for $D6^c_{\delta,\alpha}$, it can be observed that there is a visible separation between the study groups, which is confirmed by the p-values equal or lower than 0.002. In $D6^c_{\delta,\alpha}$, the groups MCI, ADM and ADA present similar mean values leading to a p-value of 0.0082, the highest value among these cases.

In addition to the global analysis, a channel analysis is performed to find evidence of activity and progression of AD in each electrode, that is, throughout the scalp. Only the features $D1^l_{\theta,\alpha}$, $D2^l_{\theta,\alpha}$, $D3^l_{\theta,\alpha}$, $D4^l_{\theta,\alpha}$, $D5^l_{\theta,\alpha}$ and $D6^c_{\delta,\alpha}$, which present the best discriminant capacity in the global analysis, are considered in this case. First, the normalized values of these distance measures are averaged only across time/segments, resulting in 19 mean normalized values of them per subject (1 per channel). Then, their mean normalized values are separated not only by AD stages (C, MCI, ADM and ADA) but also by electrode and treated as distributions.

The Kruskal-Wallis test is used on each electrode to determine if the data of the study groups (C, MCI, ADM and ADA) come from the same distribution. The significant differences, p-values lower than 0.05 after Bonferroni correction for multivariable comparisons, for All vs All and some pairs of study groups are shown in Figure 4. It can be observed that significant differences are found for all cases. The pairs C vs

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2021.3069789, IEEE Journal of Biomedical and Health Informatics

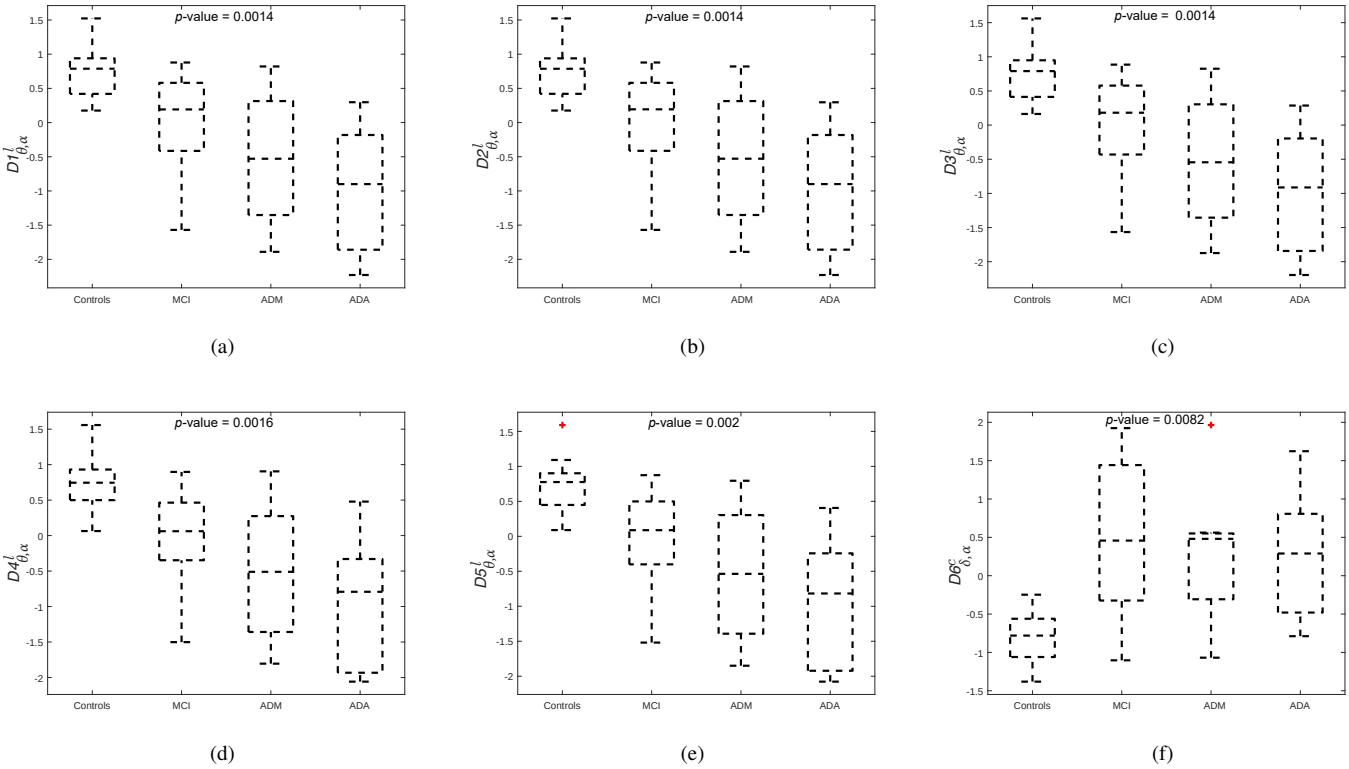


Figure 3. Boxplots of the distributions of: (a) $D1^{l}_{\theta,\alpha}$; (b) $D2^{l}_{\theta,\alpha}$; (c) $D3^{l}_{\theta,\alpha}$; (d) $D4^{l}_{\theta,\alpha}$; (e) $D5^{l}_{\theta,\alpha}$; (f) $D6^{c}_{\delta,\alpha}$.

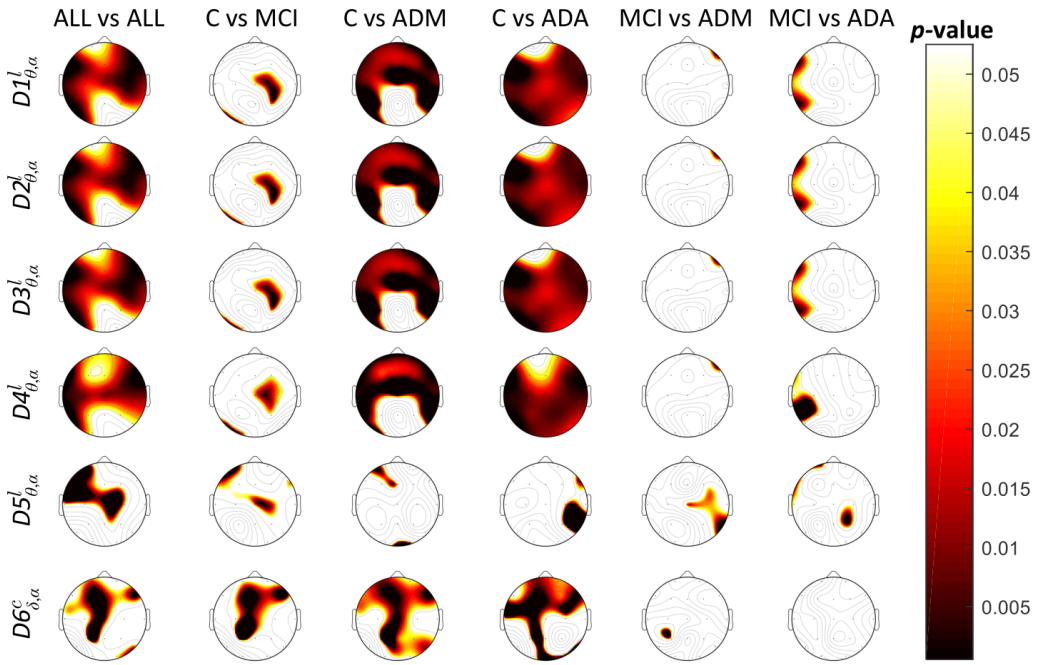


Figure 4. Sensor level topographic maps of $D1^{l}_{\theta,\alpha}$, $D2^{l}_{\theta,\alpha}$, $D3^{l}_{\theta,\alpha}$, $D4^{l}_{\theta,\alpha}$, $D5^{l}_{\theta,\alpha}$ and $D6^{c}_{\delta,\alpha}$ between all study groups (All vs All), C vs MCI, C vs ADM, C vs ADA, MCI vs ADM and MCI vs ADA. Dots indicate sensors and the statistically significant differences between groups are shown at color.

ADM and C vs ADA present the largest amount of significant differences, which indicate that the proposed features may be efficient in diagnosing AD in the stages ADM and ADA. On the other hand, the pairs MCI vs ADM and MCI vs ADA present the smallest amount of significant differences, which indicate that the proposed features may be less efficient in tracking the last two steps of AD progression. With the exception of $D6^{l}_{\delta,\alpha}$ in MCI vs ADA, all proposed features provide significant differences for all analyzed cases.

When comparing all study groups (All vs All), the sig-

nificant differences are concentrated in the parietal and right fronto-temporal regions while no significant differences are found in the occipital region. In analyzing C vs MCI, the significant differences are dense in the parietal region. For C vs ADM, the significant differences are concentrated in the front and temporal regions. In the MCI vs ADM case, the few significant differences often appear in the right fronto-temporal region. When comparing MCI vs ADA, the significant differences are concentrated in the left temporal region. In general, parietal and temporal were the most affected regions in all evaluations made. These results are in line with previous research studies where it was found that frontal, temporal and parietal regions are progressively affected as AD progresses [52], [53], [54].

B. Classification Procedure

In this section, machine learning techniques are applied to the cepstral/lacstral distances in order to develop a cepstral/lacsogram-based method for diagnosing AD in its different stages. Three different machine learning models are used as classifiers. Their configurations are described below:

- Surrogate decision trees (SDT): SDT with AdaBoost for multi-classification as train algorithm;
- Support vector machines: Multiclass SVM with radial basis function and cost factor of 10;
- Artificial neural networks: Feed-forward multilayer perceptron with tansig as activation function, Levenberg-Marquardt as the training algorithm, mean squared as the error function, output layer with 4 nodes, input layer with 11 nodes and finally one hidden layer with 30 nodes.

All classifiers are applied with a leave-one-out cross-validation. Cross-validation techniques allow the use of the entire dataset in the classification procedure without data leakage between the training and testing sets [55], [56]. It is a common procedure for getting classification conclusions in small dataset [24], [26], [29], [30], [31], [36], [57]. As can be seen in Table I, the large majority of EEG-based method for detecting AD make use of cross-validation since the availability of EEG recordings is generally small.

The leave-one-out cross-validation used in this work is depicted in Figure 5. If one features vector of a specific channel of a specific subject is selected for testing in the classification procedure, then the other 18 channel feature vectors of the same patient are automatically unused in the learning procedure, thereby ensuring data independence. A total of 722 iterations are done in the leave-one-out cross-validation, where 704 features vectors (1 for testing and 703 for training) are used in each iteration.

The classifiers are evaluated through dignostic accuracy (Accu) and receiver operating characteristic (ROC) curve. The ROC curve summarizes the performance of a multiple-class classifier and is a graphical representation of the trade-off between sensitivity (Sens) and specificity (Spec) [58]. The area under the ROC curve (AUC) is a single number that represents the diagnostic performance of the method [58].

Only the 28 significant distance measures, those that achieved a p-value lower than 0.05 in the global statistical analysis and are shown in Table III, are considered in the classification procedure. For each significant distance $Dk_{i,j}^t$, its normalized values are averaged and grouped in the same way performed in the channel analysis to be used as features. That is, the average procedure across channels is first applied, resulting in 19 mean normalized values of $Dk_{i,j}^t$ per subject (1 per channel), and then the mean normalized values are separated by AD stages (C, MCI, ADM, ADA) and channels.

In order to improve execution time and classification results, a genetic algorithm (GA) with entropy criterium [59] is applied to search the combination of 5 to 12 significant distance measures per channel that, working as input vectors (features vectors) for the classifiers, achieves the maximum accuracy for the All vs All multiclass comparison. It is noteworthy that all dataset is used for feature selection due to its scarce length. This procedure may slightly interfere in the final classification of the model but is a common procedure when the dataset is limited [24], [26], [31], [33], [35], [60].

The best features combination and the corresponding classification result for each classifier are shown in Table IV. It can be observed that the best performance of the classifiers is obtained with input vectors of different sizes, namely, 9 features for SDT, 7 features for SVM and 11 features for ANN. It can also be seen that the large majority of the combined features came from the proposed lacsogram analysis. And it can be concluded that the ANN classifier presents the best performance, achieving a sensitivity of 90.83%, a specificity of 97.73%, an accuracy of 95.55%, an AUC of 0.947 and an out-of-sample classification error of 4.73%. This result indicates that the ANN classifier generalizes fairly well. The SVM classifier outperformes the ANN one regarding specificity, obtaining 99.45%, but achieved the lowest sensitivity. As an example of the effect of the GA on classification, the best results using all significant features as input vectors are 85.53% of sensitivity, 94.56% of specificity, 90.04% of accuracy and 0.914 of AUC, achieved also with the ANN classifier.

In addition to the All vs All classification, classifications between pairs of study groups are also performed. Thereunto, the same ANN architecture, the same method of leave-one-out cross-validation and the same 11 features selected by GA for the All vs All classification are used. The classification results for each pair of study groups are shown in Table V. It can be observed that the lowest classification accuracy is 88.84% for MCI vs ADM. This may be due to the small amount of significant differences at the scalp level for this specific comparison as indicated in Figure 4. On the other hand, in all other pairs of study groups, the proposed method achieves classification accuracies greater than 94%. For C vs MCI and MCI vs ADA, it achieves accuracies greater than 98%.

A comparison between the classification results found in the present work and in the literature is shown in Table VI where, to simplify, each result is rounded to the nearest integer number. It can be observed that most state-of-art methods have focused on discriminating some pairs of groups, mainly C vs ADM, C vs MCI and MCI vs ADM-ADA, but only few have proposed to distinguish all AD stages (All vs All).

Compared with methods of diagnosing AD through biomarkers others than EEG, it can be observed that the pro-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2021.3069789, IEEE Journal of Biomedical and Health Informatics

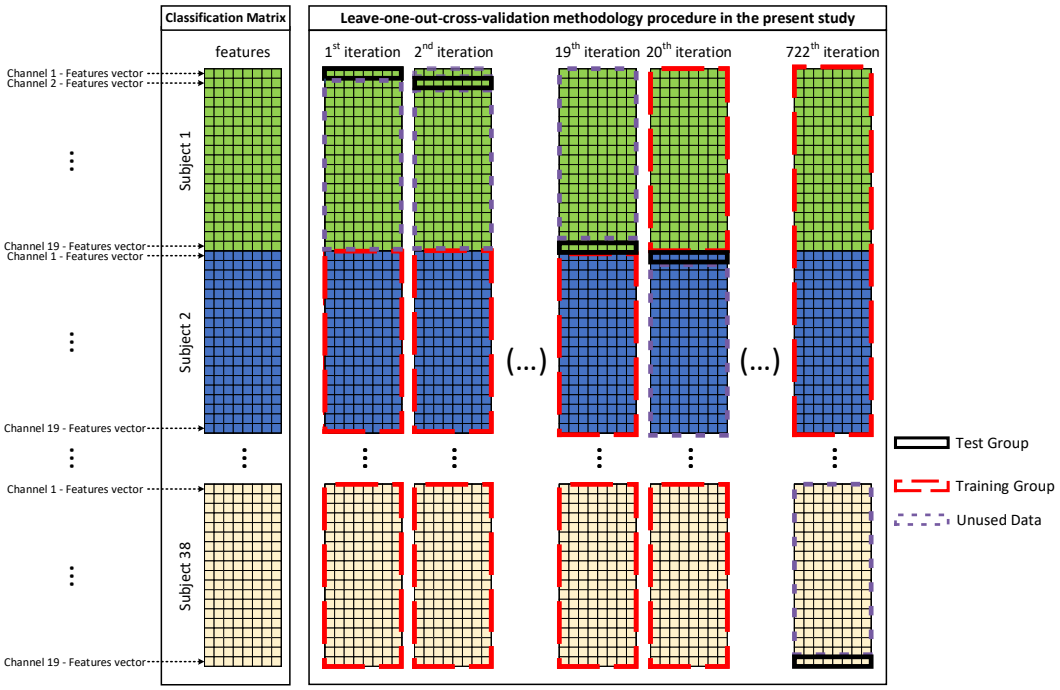


Figure 5. Graphical illustration of the leave-one-out cross-validation used in the present work.

Table IV
BEST CLASSIFICATION RESULTS WITH DIFFERENT KIND OF CLASSIFIERS FOR ALL VS ALL.

Classifier	# of features	Features Combination	Sens (%)	Spec (%)	Accu (%)	AUC
SDT	9	$D4^{l}_{\theta,\alpha}, D1^{l}_{\theta,\alpha}, D2^{l}_{\theta,\alpha}, D1^{l}_{\theta,\beta}, D6^{c}_{\delta,\alpha}, D1^{l}_{\delta,\alpha}, D6^{l}_{\alpha,\gamma}, D4^{l}_{\theta,\beta}, D2^{l}_{\theta,\alpha}$	88.79	97.71	94.88	0.936
SVM	7	$D4^{l}_{\theta,\alpha}, D1^{l}_{\theta,\alpha}, D2^{l}_{\theta,\alpha}, D1^{l}_{\theta,\beta}, D6^{c}_{\delta,\alpha}, D1^{l}_{\delta,\alpha}, D6^{l}_{\alpha,\gamma}$	87.24	99.45	95.10	0.937
ANN	11	$D4^{l}_{\theta,\alpha}, D1^{l}_{\theta,\alpha}, D2^{l}_{\theta,\alpha}, D1^{l}_{\theta,\beta}, D6^{c}_{\delta,\alpha}, D1^{l}_{\delta,\alpha}, D6^{l}_{\alpha,\gamma}, D4^{l}_{\theta,\beta}, D2^{l}_{\theta,\alpha}, D1^{c}_{\delta,\alpha}, D5^{l}_{\delta,\alpha}$	90.83	97.73	95.55	0.947

Table V
CLASSIFICATION RESULTS WITH ANN CLASSIFIER FOR PAIRS OF STUDY GROUPS.

C vs MCI			C vs ADM			C vs ADA			C vs ADM-ADA			MCI vs ADM			MCI vs ADA			MCI vs ADM-ADA		
Sens (%)	**Spec** (%)	**Accu** (%)	**Sens** (%)	**Spec** (%)	**Accu** (%)	**Sens** (%)	**Spec** (%)	**Accu** (%)	**Sens** (%)	**Spec** (%)	**Accu** (%)	**Sens** (%)	**Spec** (%)	**Accu** (%)	**Sens** (%)	**Spec** (%)	**Accu** (%)	**Sens** (%)	**Spec** (%)	**Accu** (%)
99.03	96.77	98.06	96.17	95.79	95.99	97.09	94.19	95.84	96.63	94.99	95.92	85.00	88.89	86.84	98.44	99.33	98.83	91.72	96.34	93.85

posed method achieves the highest classification accuracies. In the All vs All and C vs MCI cases, it outperformes the method based on cerebrospinal biomarkers and image developed in Palmqvist *et al.* [13] by 2% and 8%, respectively. For the pair MCI vs ADM-ADA, the proposed method outperformes the MMSE-based method introduced in O'Bryant *et al.* [63] by impressive 11%. For C vs ADM, it outperformes the CSF-based method proposed in Shaw *et al.* [10] by 9%.

With regard to AD diagnosis through EEG signals, it can also be noticed that the proposed method achieves the highest classification accuracies. In the All vs All case, the proposed method outperformes the methods developed in Acrofi *et al.* [24], Huang *et al.* [30] and Rodrigues *et al.* [37] by 19%, 18% and 6%, respectively. For the pair C vs MCI, it outperformes the EEG bump modulation method introduced in Vialatte *et al.* [36] by 5%. In MCI vs ADM-ADA, the proposed method outperformes the IFAST method developed in Buscema *et al.* [28] by 2%. For C vs ADM, it outperformes the DWT-based method proposed in Rodrigues *et al.* [37] by 1%. It should be noted that, except for [37], the EEG databases of the aforementioned works differ from that of the present one.

Finally, a comparison between the present work and the previous ones with the same EEG database is presented in Table VII. As can be seen, the present work is the only one to use the multiband lacsogram analysis and the corresponding lactral distances $Dk^{l}_{i,j}$. Due to the high discriminant capacity of theses features, the present work represents an improvement in both statistical (p-value) and classification results.

Therefore, as it achieves high classification accuracies ($>$86%) for all study comparisons and classification accuracies higher than the state-of-art methods, it can be concluded that the proposed method represents an advance in diagnosing AD in all its stages. The improvements achieved in C vs MCI, C vs ADM and MCI vs ADM-ADA indicates that this advance is significant in early detection which, due to the lack of a cure, plays a key role in medical intervention to reduce brain damage, preserve daily functioning for longer and give the patient time to plan the future.

Table VI
COMPARISON BETWEEN THE CLASSIFICATION RESULTS OF PRESENT STUDY AND STATE-OF-ART.

Exam Types	Authors	C vs ADM			C vs MCI			MCI vs ADM-ADA			All vs All		
		Sens (%)	Spec (%)	Accu (%)	Sens (%)	Spec (%)	Accu (%)	Sens (%)	Spec (%)	Accu (%)	Sens (%)	Spec (%)	Accu (%)
EEG	Akrofi *et al.* [24]	-	-	-	-	-	90	-	-	-	-	-	77
	Afshari *et al.* [25]	-	-	94	-	-	-	-	-	-	-	-	-
	Aghajani *et al.* [26]	-	-	-	75	94	84	-	-	-	-	-	-
	Besthorn *et al.* [27]	-	-	86	-	-	-	-	-	-	-	-	-
	Buscema *et al.* [28]	-	-	-	-	-	-	89	95	92	-	-	-
	Cassani *et al.* [29]	-	-	79	-	-	-	-	-	-	-	-	-
	Huang *et al.* [30]	90	75	84	-	-	-	87	68	78	87	68	78
	Khatun *et al.* [31]	-	-	-	85	95	88	-	-	-	-	-	-
	Knott *et al.* [32]	-	-	75	-	-	-	-	-	-	-	-	-
	Melissant *et al.* [33]	93	95	94	-	-	-	-	-	-	-	-	-
	Petrosian *et al.* [34]	80	100	90	-	-	-	-	-	-	-	-	-
	Poil *et al.* [35]	-	-	-	-	-	-	88	82	85	-	-	-
	Viallate *et al.* [36]	-	-	-	-	-	93	-	-	-	-	-	-
	Rodrigues *et al.* [37]	96	93	95	80	75	77	84	82	83	87	94	90
	Present work	**97**	**95**	**96**	**100**	**97**	**98**	**92**	**96**	**94**	**91**	**98**	**96**
Blood & cerebrospinal Biomarkers + Image Methods	Colliot *et al.* [14]	84	84	84	-	-	-	-	-	-	-	-	-
	Forlenza *et al.* [61]	88	78	83	-	-	-	90	65	78	-	-	-
	Hampel *et al.* [12]	80	90	85	-	-	-	-	-	-	-	-	-
	Li *et al.* [57]	92	79	85	-	-	-	-	-	-	-	-	-
	Palmqvist *et al.* [13]	85	75	80	80	86	83	-	-	-	-	-	-
	Palmqvist *et al.* [13]	56	93	75	97	83	90	-	-	-	-	-	94
	Shaw *et al.* [10]	96	77	87	-	-	-	-	-	-	-	-	-
MMSE	Freitas *et al.* [62]	-	-	-	65	97	81	-	-	-	-	-	-
	Knopman *et al.* [20]	81	70	76	-	-	-	-	-	-	-	-	-
	O'Bryant *et al.* [63]	-	-	-	-	-	-	66	99	83	-	-	-

Table VII
COMPARISON BETWEEN PRESENT STUDY AND PREVIOUS WORKS WITH THE SAME EEG DATABASE.

Ref.	Summary	EEG Signal Processing	Extracted features	Statistical analysis results (Kruskal-Wallis with multi-variable correction)	Classification accuracy				
					C vs MCI	C vs AD	MCI vs AD	All vs All	Best Classifier
[38]	A metric of event propagation time was used to differentiate C, MCI and ADM. Its discrimination capacity was analyzed through Kruskal-Wallis test. Results suggested that there are energy variation sequences that appear more frequently in ADM. The average of an event propagation time in ADM, MCI and C is approximately 850, 790 and 680 ms, respectively.	Multiband spectral analysis via DWT	Event propagation time (ETEP)	C vs MCI vs ADM: ETEP with a p-value=0.025	No classification procedure was applied				
[39]	A cepstral multiband analysis was applied to discriminate between C, MCI, ADM and ADA. The cepstral distances $Dk^c_{i,j}$ were computed and then statistically analyzed through the Kruskal-Wallis test.	Multiband cepstral analysis via DWT	Cepstral distances $Dk^c_{i,j}$	C vs MCI vs ADM vs ADA: $D2^c_{\theta,\alpha}$ and $D3^c_{\theta,\alpha}$ with p-value=0.006	No classification procedure was applied				
[37]	A multiband spectral analysis was applied to extract features to differentiate C, MCI, ADM and ADA. The individual discrimination capacity of the features was analyzed through Kruskal-Wallis test and the relative power in subband β presented the best result. A set of machine learning tools were fed with the significant features for classification.	Multiband spectral analysis via DWT	Conventional subbands relative power (RP), spectral ratios, number of maxima, minima and zero crossing	C vs MCI vs ADM vs ADA: RP_β with p-value=0.012	77%	95%	83%	90%	ANN
Present work	A multiband cepstral and lacstral analysis were applied to extract features to discriminate between C, MCI, ADM and ADA. The cepstral and lacstral distances $Dk^t_{i,j}$ were computed and then statistically analyzed through the Kruskal-Wallis test. The lacstral distances $D1^l_{\theta,\alpha}$, $D2^l_{\theta,\alpha}$ and $D3^l_{\theta,\alpha}$ presented the best result. A set of machine learning tools were fed with the best features combination provided by a genetic algorithm for classification.	Multiband cepstral and lacstral analysis via DWT	Cepstral $Dk^c_{i,j}$ and lacstral $Dk^l_{i,j}$ distances	C vs MCI vs ADM vs ADA: $D1^l_{\theta,\alpha}$, $D2^l_{\theta,\alpha}$ and $D3^l_{\theta,\alpha}$ with a p-value=0.0014	98%	96%	94%	96%	ANN

V. CONCLUSION

This work proposed a new signal tool, called lacsogram, for application on EEG signals to characterize and diagnose Alzheimer's disease in different stages: Controls, Mild Cognitive Impairment, Mild and Moderate AD and Advanced AD.

Statistical analyzes were performed to several lacstral distances between conventional EEG subbands to find measures capable of discriminating AD in all stages as well as characterizing activity and progress of AD in each electrode. Cepstral distances were used for comparison. The results showed that the lacsogram provided a greater amount of significant distances, the 15 most significant distances and the highest significance value for 5 of the 6 distance measures. At sensor level, the topographic maps showed significant differences in parietal, temporal and frontal regions as AD progresses.

A genetic algorithm was used to select significant distance measures as inputs for different classifiers with leave-one-out cross-validation to diagnose AD in different stages. The ANN classifier presented the best performance and its classification accuracies were: 95.55% for All vs All, 98.06% for C vs MCI, 95.99% for C vs ADM, 93.85% for MCI vs ADM-ADA.

In C vs MCI, C vs ADM and MCI vs ADM-ADA, the proposed method outperformed the state-of-art methods by 5%, 1% and 2%, respectively. In All vs All, it outperformed the state-of-art EEG and non-EEG methods by 6% and 2%, respectively. Therefore, the proposed method represents an advance in diagnosing AD, mainly in early stages.

Despite the promising results, which indicate a great ability of the proposed method to diagnose AD in its different stages, the database used in this work is slightly limited. In the future, the results should be updated with a larger population to ensure a consistent generalization and an analysis of whether the classifier choice significantly affects the results should be done.

ACKNOWLEDGMENT

The authors would like to thank the Neurological Unity of University Hospital Center of São João, Porto, Portugal, for supplying the EEG signals. This work was supported by National Funds from FCT – Fundação para a Ciência e a Tecnologia through projects UIDB/50016/2020 and UIDB/05757/2020 and by the Doctoral Program in Biomedical Engineering (PRODEB) of the Faculty of Engineering of the University of Porto.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

[1] C. Ballard, S. Gauthier, A. Corbett, C. Brayne, D. Aarsland, and E. Jones, "Alzheimer's disease," *The Lancet*, vol. 377, pp. 1019–1031, March 2011.

[2] Z. Song, B. Deng, J. Wang, and R. Wang, "Biomarkers for Alzheimer's disease defined by a novel brain functional network measure," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 1, pp. 41–49, Jan. 2019.

[3] World Alzheimer Report 2018, "The state of the art of dementia research: new frontiers," Alzheimer's Disease International, London, UK, September 2018.

[4] T. Bird, *Harrison's Principles of Internal Medicine*, 15th ed. McGraw-Hill, 2001, ch. Alzheimer's disease and other primary dementias, pp. 2391–2399.

[5] P. Nestor, P. Scheltens, and J. Hodges, "Advances in the early detection of Alzheimer's disease," *Nat Med*, vol. 10, pp. S34–S41, July 2004.

[6] T. Tong, Q. Gao, R. Guerrero, C. Ledig, L. Chen, D. Rueckert, and A. D. N. Initiative, "A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 155–165, Jan. 2017.

[7] M. M. Mesulam, *Principies of behavioural and cognitive neurology*, 2nd ed. New York, NY: Oxford University Press, 2000, ch. Aging Alzheimer disease and dementia, pp. 439–522.

[8] J. Jeong, "EEG dynamics in patients with Alzheimer's disease," *Clinical Neurophysiology*, vol. 115, no. 7, pp. 1490–1505, July 2004.

[9] K. Blennow, "PL.02.01 CSF biomarkers in Alzheimer's disease - use in clinical diagnosis and to monitor treatment effects," *European Neuropsychopharmacology*, vol. 20, p. S159, Aug. 2010.

[10] L. M. Shaw, H. Vanderstichele, M. Knapik-Czajka, C. M. Clark, P. S. Aisen, R. C. Petersen, K. Blennow, H. Soares, A. Simon, P. Lewczuk, R. Dean, E. Siemers, W. Potter, V. M.-Y. Lee, and J. Q. Trojanowski, "Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects," *Annals of Neurology*, vol. 65, no. 4, pp. 403–413, April 2009.

[11] G. De Meyer, F. Shapiro, H. Vanderstichele, E. Vanmechelen, S. Engelborghs, P. P. De Deyn, E. Coart, O. Hansson, L. Minthon, H. Zetterberg, K. Blennow, L. Shaw, and J. Q. Trojanowski, "Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people," *Archives of Neurology*, vol. 67, no. 8, pp. 949–956, Aug. 2010.

[12] H. Hampel, R. Frank, K. Broich, S. Teipel, R. Katz, and J. Hardy, "Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives," *Nat Rev Drug Discovery*, vol. 9, pp. 560–574, July 2010.

[13] S. Palmqvist, H. Zetterberg, N. Mattsson, P. Johansson, L. Minthon, K. Blennow, M. Olsson, and O. Hansson, "Detailed comparison of amyloid PET and CSF biomarkers for identifying early Alzheimer disease," *Neurology*, vol. 85, no. 14, pp. 1240–1249, Oct. 2015.

[14] O. Colliot, G. Chetelat, M. Chupin, B. Desgranges, B. Magnin, H. Benali, B. Dubois, L. Garnero, F. Eustache, and S. Lehericy, "Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus," *Radiology*, vol. 248, no. 1, pp. 194–201, July 2008.

[15] H. Barthel, H.-J. Gertz, S. Dresel, O. Peters, P. Bartenstein, K. Buerger, F. Hiemeyer, S. M. Wittemer-Rump, J. Seibyl, C. Reininger, and O. Sabri, "Cerebral amyloid-β PET with florbetaben (18F) in patients with Alzheimer's disease and healthy controls: a multicentre phase 2 diagnostic study," *Lancet Neurology*, vol. 10, no. 5, pp. 424–435, May 2011.

[16] N. Tolboom, W. M. van der Flier, J. Boverhoff, M. Yaqub, M. P. Wattjes, P. G. Raijmakers, F. Barkhof, P. Scheltens, K. Herholz, A. A. Lammertsma, and B. N. M. van Berckel, "Molecular imaging in the diagnosis of Alzheimer's disease: visual assessment of [11C]PIB and [18F]FDDNP PET images," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 81, no. 8, pp. 882–884, Aug. 2010.

[17] K. Herholz, "PET studies in dementia," *Annals of Nuclear Medicine*, vol. 17, no. 2, pp. 79–89, April 2003.

[18] G. M. McKhann, D. S. Knopman, H. Chertkow, B. T. Hyman, C. R. Jack Jr, C. H. Kawas, W. E. Klunk, W. J. Koroshetz, J. J. Manly, R. Mayeux, R. C. Mohs, J. C. Morris, M. N. Rossor, P. Scheltens, M. C. Carrillo, B. Thies, S. Weintraub, and C. H. Phelps, "The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's & Dementia*, vol. 7, no. 3, pp. 263–269, May 2011.

[19] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease," *Neurology*, vol. 34, no. 7, pp. 939–944, July 1984.

[20] D. S. Knopman, S. T. DeKosky, J. L. Cummings, H. Chui, J. Corey–Bloom, N. Relkin, G. W. Small, B. Miller, and J. C. Stevens, "Practice parameter: diagnosis of dementia (an evidence-based review)," *Neurology*, vol. 56, no. 9, pp. 1143–1153, May 2001.

[21] J. Hort, J. T. O'Brien, G. Gainotti, T. Pirttila, B. O. Popescu, I. Rektorova, S. Sorbi, and P. Scheltens, "EFNS guidelines for the diagnosis and management of Alzheimer's disease," *European Journal of Neurology*, vol. 17, no. 10, pp. 1236–1248, Oct. 2010.

[22] S. Sanei and J. Chambers, *EEG signal processing*. Chichester, England: John Wiley & Sons, 2007.

[23] J. Dauwels, F.-B. Vialatte, and A. Cichocki, "On the early diagnosis of Alzheimer's disease from EEG signals: a mini-review," in *Advances in Cognitive Neurodynamics (II)*, R. Wang and F. Gu, Eds. Dordrecht: Springer, Oct. 2010, pp. 709–716.

[24] K. Akrofi, R. Pal, M. C. Baker, B. S. Nutter, and R. W. Schiffer, "Classification of Alzheimer's disease and mild cognitive impairment by pattern recognition of EEG power and coherence," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, USA, March 2010, pp. 606–609.

[25] S. Afshari and M. Jalili, "Directed functional networks in Alzheimer's disease: disruption of global and local connectivity measures," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 4, pp. 949–955, July 2017.

[26] H. Aghajani, E. Zahedi, M. Jalili, A. Keikhosravi, and B. V. Vahdat, "Diagnosis of early Alzheimer's disease based on EEG source localization and a standardized realistic head model," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 6, pp. 1039–1045, Nov. 2013.

[27] C. Besthorn, H. Förstl, C. Geiger-Kabisch, H. Sattel, T. Gasser, and U. Schreiter-Gasser, "EEG coherence in Alzheimer's disease," *Electroencephalography and Clinical Neurophysiology*, vol. 90, no. 3, pp. 242–245, March 1994.

[28] M. Buscema, P. Rossini, C. Babiloni, and E. Grossi, "The IFAST model, a novel parallel nonlinear EEG analysis technique, distinguishes mild cognitive impairment and Alzheimer's disease patients with high degree of accuracy," *Artificial Intelligence in Medicine*, vol. 40, no. 2, pp. 127–141, June 2007.

[29] R. Cassani and T. Falk, "Alzheimer's disease diagnosis and severity level detection based on electroencephalography modulation spectral "patch" features," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 1–12, July 2019.

[30] C. Huang, L.-O. Wahlund, T. Dierks, P. Julin, B. Winblad, and V. Jelic, "Discrimination of Alzheimer's disease and mild cognitive impairment by equivalent EEG sources: a cross-sectional and longitudinal study," *Clinical Neurophysiology*, vol. 111, no. 11, pp. 1961–1967, Nov. 2000.

[31] S. Khatun, B. I. Morshed, and G. M. Bidelman, "A single-channel EEG-based approach to detect mild cognitive impairment via speech-evoked brain responses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 5, pp. 1063–1070, May 2019.

[32] V. Knott, E. Mohr, C. Mahoney, and V. Ilivitsky, "Quantitative electroencephalography in Alzheimer's disease: comparison with a control group, population norms and mental status," *Journal of Psychiatry & Neuroscience*, vol. 26, no. 2, pp. 106–116, March 2001.

[33] C. Melissant, A. Ypma, E. E. E. Frietman, and C. J. Stam, "A method for detection of Alzheimer's disease using ICA-enhanced EEG measurements," *Artificial Intelligence in Medicine*, vol. 33, no. 3, pp. 209–222, March 2005.

[34] A. A. Petrosian, D. V. Prokhorov, W. Lajara-Nanson, and R. B. Schiffer, "Recurrent neural network-based approach for early recognition of Alzheimer's disease in EEG," *Clinical Neurophysiology*, vol. 112, no. 2, pp. 1378–1387, Aug. 2001.

[35] S.-S. Poil, W. de Haan, W. M. van der Flier, H. D. Mansvelder, P. Scheltens, and K. Linkenkaer-Hansen, "Integrative EEG biomarkers predict progression to Alzheimer's disease at the MCI stage," *Frontiers in Aging Neuroscience*, vol. 5, p. 58, Oct. 2013.

[36] F. Vialatte, A. Cichocki, G. Dreyfus, T. Musha, S. L. Shishkin, and R. Gervais, "Early detection of Alzheimer's disease by blind source separation, time frequency representation, and bump modeling of EEG signals," in *Artificial Neural Networks: Biological Inspirations - ICANN 2005*, ser. Lecture Notes in Computer Science, W. Duch, J. Kacprzyk, E. Oja, and S. Zadrożny, Eds., vol. 3696. Berlin, Heidelberg: Springer, 2005, pp. 683–692.

[37] P. M. Rodrigues, D. R. Freitas, J. P. Teixeira, D. Alves, and C. Garrett, "Electroencephalogram signal analysis in Alzheimer's disease early detection," *International Journal of Reliable and Quality E-Healthcare*, vol. 7, no. 1, pp. 40–59, Jan. 2018.

[38] P. M. Rodrigues, D. Freitas, J. P. Teixeira, D. Alves, and C. Garrett, "Early detection of electroencephalogram temporal events in Alzheimer's disease," in *Design, Development, and Integration of Reliable Electronic Healthcare Platforms*, A. Moumtzoglou, Ed. Hershey, PA: IGI Global, 2017, pp. 112–131.

[39] P. M. Rodrigues, D. Freitas, J. P. Teixeira, B. Bispo, D. Alves, and C. Garrett, "Electroencephalogram hybrid method for Alzheimer early detection," *Procedia Computer Science*, vol. 138, pp. 209–214, 2018.

[40] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[41] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, MA: Artech House, 1992.

[42] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14–38, Oct. 1992.

[43] B. Bogert, M. Healy, and J. Tukey, *Time Series Analysis*. Wiley, 1963, ch. The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking, pp. 209–243.

[44] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley - Cambridge Press, 1996.

[45] A. Gray and J. Markel, "Distance measures for speech processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 5, pp. 380–391, Oct. 1976.

[46] Y. Tohkura, "A weighted cepstral distance measure for speech recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 10, pp. 1414–1422, Oct. 1987.

[47] K. Paliwal, "On the performance of the quefrency-weighted cepstral coefficients in vowel recognition," *Speech Communication*, vol. 1, no. 2, pp. 151–154, Jan. 1982.

[48] K. L. Priddy and P. E. Keller, *Artificial Neural Networks: An Introduction.* Bellingham, WA: SPIE Press, 2005.

[49] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 583–621, Dec. 1952.

[50] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, Sep. 1979.

[51] A. Agnoli, N. Martucci, V. Manna, L. Conti, and M. Fioravanti, "Effect of cholinergic and anticholinergic drugs on short-term memory in Alzheimer's dementia: a neuropsychological and computerized electroencephalographic study." *Clinical Neuropharmacology*, vol. 6, no. 4, pp. 311–324, Dec. 1983.

[52] H. Hampel, D. Prvulovic, S. Teipel, F. Jessen, C. Luckhaus, L. Frölich, M. W. Riepe, R. Dodel, T. Leyhe, L. Bertram, W. Hoffmann, and F. Faltraco, "The future of Alzheimer's disease: the next 10 years," *Progress in Neurobiology*, vol. 95, no. 4, pp. 718–728, Dec. 2011.

[53] R. A. Sperling, P. S. Aisen, L. A. Beckett, D. A. Bennett, S. Craft, A. M. Fagan, T. Iwatsubo, C. R. Jack, J. Kaye, T. J. Montine, D. C. Park, E. M. Reiman, C. C. Rowe, E. Siemers, Y. Stern, K. Yaffe, M. C. Carrillo, B. Thies, M. Morrison-Bogorad, M. V. Wagster, and C. H. Phelps, "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*, vol. 7, no. 3, pp. 280–292, May 2011.

[54] I. Ferrer, "Defining Alzheimer as a common age-related neurodegenerative process not inevitably leading to dementia," *Progress in Neurobiology*, vol. 97, no. 1, pp. 38–51, Apr. 2012.

[55] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ: Pearson, 2009.

[56] T. Trappenberg, *Fundamentals of Machine Learning*. New York, NY: Oxford University Press, 2019.

[57] W. Li, Y. Zhao, X. Chen, Y. Xiao, and Y. Qin, "Detecting Alzheimer's disease on small dataset: a knowledge transfer perspective," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1234–1242, May 2019.

[58] C. J. Williams, S. S. Lee, R. A. Fisher, and L. H. Dickerman, "A comparison of statistical methods for prenatal screening for Down syndrome," *Applied Stochastic Models in Business and Industry*, vol. 15, no. 2, pp. 89–101, April 1999.

[59] S. K. Smit and A. E. Eiben, "Using entropy for parameter analysis of evolutionary algorithms," in *Experimental Methods for the Analysis of Optimization Algorithms*, T. Bartz-Beielstein, M. Chiarandini, L. Paquete, and M. Preuss, Eds. Berlin, Heidelberg: Springer, 2010, pp. 287–310.

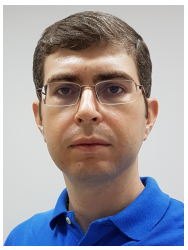
[60] E. Gallego-Jutglà, J. Solé-Casals, F.-B. Vialatte, M. Elgendi, A. Cichocki, and J. Dauwels, "A hybrid feature selection approach for the early diagnosis of Alzheimer's disease," *Journal of Neural Engineering*, vol. 12, no. 1, p. 016018, Feb. 2015.

[61] O. V. Forlenza, M. Radanovic, L. L. Talib, I. Aprahamian, B. S. Diniz, H. Zetterberg, and W. F. Gattaz, "Cerebrospinal fluid biomarkers in Alzheimer's disease: diagnostic accuracy and prediction of dementia," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 1, no. 4, pp. 455–463, Dec. 2015.

[62] S. Freitas, M. R. Simões, L. Alves, and I. Santana, "Montreal Cognitive Assessment (MoCA): normative study for the Portuguese population," *Journal of Clinical and Experimental Neuropsychology*, vol. 33, no. 9, pp. 989–996, Nov. 2011.

[63] S. E. O'Bryant, J. D. Humphreys, G. E. Smith, R. J. Ivnik, N. R. Graff-Radford, R. C. Petersen, and J. A. Lucas, "Detecting dementia with the mini-mental state examination in highly educated individuals," *Archives of Neurology*, vol. 65, no. 7, pp. 963–967, July 2008.

Pedro Miguel Rodrigues received the B.Sc. and M.Sc. degrees in biomedical engineering from the Polytechnic Institute of Bragança, Portugal, in 2009 and 2011, respectively, and the Ph.D. degree in biomedical engineering from the University of Porto, Portugal, in 2017. Since 2018, he is an Assistant Professor at the Faculty of Biotechnology – Universidade Católica Portuguesa (ESB-UCP) and a full member of the research Centre for Biotechnology and Fine Chemistry (CBQF-UCP), Porto, Portugal. His research areas focus on digital signal processing and artificial intelligence applied to the diagnosis of neurodegenerative, cardiac and speech diseases.



Bruno Catarino Bispo received the B.Sc. degree in electronics and computer engineering from the Federal University of Rio de Janeiro (UFRJ), Brazil, in 2006, the M.Sc. degree in electrical engineering from UFRJ in 2008, and the Ph.D. degree in electrical and computer engineering from the University of Porto, Portugal, in 2015. From 2009 to 2012, he was a Lecturer at the Faculty of Engineering of the University of Porto. From 2015 to 2019, he was an Assistant Professor at the Federal University of Technology – Paraná (UTFPR), Brazil. Since 2019, he has been an Assistant Professor in the Department of Electrical and Electronic Engineering at the Federal University of Santa Catarina (UFSC), Florianópolis, Brazil. His research interests are in digital signal processing, particularly speech and biomedical signal processing and adaptive systems.

Carolina Garrett obtained her graduation in Medicine in 1975 at Faculty of Medicine of Porto University, Porto, Portugal. She is neurologist since 1986 and got her PhD in Neurology in 1992. She is associated Professor of Neurology, Head of the Department of Neurology in Centro Hospitalar de São João and Head of the unit of Neurology and Neurosurgery of the Department in Clinical Neurosciences and Mental health of the Faculty of Medicine Porto University.

Dílio Alves obtained his graduation in Medicine at Faculdade de Medicina da Universidade do Porto in 1974. In January of 1984 he specialized in Neurology. Since 2011, he is responsible for the Neurophysiology Unit of the Neurology Service of Hospital São João - Porto and also Coordinator of the Reference Center for Refractory Epilepsy since 2016.

João Paulo Texeira holds a PhD in Electrical and Computers Engineering from the University of Porto (2004) and is Coordinator Professor at Instituto Politécnico de Bragança – Portugal, since 1995. He teaches in the areas of Signal Processing, Electronics and Computers Engineering. He is full member of Research Centre in Digitalization and Intelligent Robotics (CeDRI). Published 2 scientific books, more than 100 scientific journal/conference papers and belongs to the Scientific Committee of several journals. Currently, his research interest focuses on AI applied to biological signals, namely EEG for the diagnosis of AD, ECG, EMG and voice, speech signals (analysis, synthesis and prosody) and time series forecasting.

Diamantino Rui da Silva Freitas graduated in 1976 and obtained a PhD in Electrical and Computer from the University of Porto in 1991. He has held teaching functions in the Electrical Engineering Department of the Faculty of Engineering of the University of Porto since 1975, being currently an associate professor and coordinator of the Speech Processing, Electro-acoustics, Signals and Instrumentation Laboratory (LPF-ESI). Participated in several R&D actions, ocasionally as national delegate, focusing on the areas of electronic instrumentation, speech processing and biomedical and rehabilitation engineering. Authored or co-authored more than 180 papers in international conferences, international book chapters and international journals, and was awarded with 2 patents.