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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

Alzheimer's Disease Diagnosis and Severity Level Detection Based on Electroencephalography Modulation Spectral "Patch" Features

Raymundo Cassani, *Member, IEEE,* Tiago H. Falk, *Senior member, IEEE*

***Abstract*—Over the last two decades, electroencephalography (EEG) has emerged as a reliable tool for the diagnosis of cortical disorders such as Alzheimer's disease (AD). Typically, resting-state EEG (rsEEG) signals have been used, and traditional frequency bands (delta, theta, alpha, beta and gamma) have been explored. Recent studies, however, have suggested that non-conventional bands may lead to improved diagnostic performance. In this work, we propose a new type of features derived from the 2-dimensional modulation spectral domain representation of the rsEEG signal in order to characterize the neuromodulatory deficit emergent with AD. The proposed features are computed as the power in specific "patches" or regions of interest in the power modulation spectrogram, which are shown to be highly discriminant of AD severity levels. The proposed features were compared with traditional features used in the rsEEG AD monitoring literature. Results showed the proposed features not only achieving improved performance at discriminating between healthy normal elderly controls (Nold) and AD patients with varying severity levels, but also at monitoring severity levels (i.e., mild AD versus moderate AD). Moreover, the proposed features were shown to outperform traditional rsEEG features. Finally, we validated the biological origin of the proposed features by using source localization and comparing the obtained results with ones reported in the AD literature.**

***Index Terms*—Alzheimer's disease, amplitude modulation, electroencephalography, resting-state, source localization.**

I. Introduction

ALZHEIMER'S disease (AD) is a progressive neurodegenerative disorder that accounts for nearly 70% of dementia cases worldwide. Projections estimate that by 2030, 75 million cases will exist worldwide [1]. While there is no cure for AD, early diagnosis can improve the quality-of-life of patients, their families and caregivers. Moreover, as novel disease-modifying drugs are being studied and developed, it is likely that their efficacy would be higher if administered during the early stages of the disease [2]. Today AD diagnosis is carried out using medical records and exhaustive mental status examinations. For symptomatic individuals, diagnosis is supported by biomarkers derived from cerebrospinal fluid (CSF) and neuroimaging techniques such as magnetic resonance imaging (MRI), computed tomography (CT) and positron emission tomography (PET) [3]. The invasive methods required to obtain CSF samples, along with the restrictive cost and lack of portability of neuroimaging techniques, have, however, hindered their use in daily clinical practice, as well as in low-income countries and in remote regions. As such, biomarkers based on electroencephalography (EEG) have emerged as a promising tool in the study of AD.

EEG signals have their origin in the underlying activity in the cerebral cortex, thus EEG-based biomarkers can be used to infer neuronal degeneration and decay in the number of synapses caused by AD progression [4]. Most of the published works have relied on the analysis of resting-state EEG (rsEEG) recordings, also known as spontaneous or background EEG. With rsEEG, participants are not required to perform any specific task, thus making recordings simpler and more comfortable for elderly participants [5], [6], [7].

Over the last two decades, several works have repeatedly reported three major effects of AD and its progression on rsEEG signals, namely, *slowing*, *decrease in synchronization* and *reduced complexity* [8]. Moreover, recent experimental evidence [9], [10] has suggested a *neuromodulatory deficit* in the rsEEG signals, which may have its origin in the reduction of neurotransmitters due to damage in brain pathways caused by AD [11], [12]. To characterize these neuromodulatory deficits, amplitude modulation (AM) analysis of rsEEG signals was recently proposed [13], [14]. While these four effects are often regarded independently, they all originate from the loss of neurons that disrupt anatomical brain connectivity at the level of functional networks [15], [16], [11]. Given these insights, several rsEEG biomarkers (or features) have been proposed in the literature to either diagnose AD to monitor disease progression [17], [18]. While the literature has reported accuracy levels around 70-90% [11], [5], [19], due to differences in testing setups, data collection protocols, and database sizes, it is hard to compare studies and determine a *de facto* state-of-the-art system. Notwithstanding, the interested reader is referred to the following comprehensive reviews for a complete view of existing EEG features for AD characterization [11], [20], [8].

Typically, rsEEG signals are analyzed in the five traditional frequency bands: delta (0.1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz) and gamma (>30 Hz). However, these bands have been defined based on visual inspection of rsEEG signals in healthy adults [4], thus they may not be optimal for different conditions, in this sense, non-traditional frequency bands have been proven useful to study EEG in neonatal period and childhood [21], and schizophrenia in adults [22]. Regarding the rsEEG-based study of AD, recent works, [23] and [24], investigated the discrimination power of EEG fea-

R. Cassani and T.H. Falk are with the Institut National de la Recherche Scientifique (INRS-EMT), University of Quebec, Montreal, QC, Canada e-mail: (see http://musaelab.ca/).

Manuscript received XXXXX; revised XXXXX.

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

tures extracted from non-traditional frequency bands. In the work presented in [23], the lower and upper frequency bounds for different power spectral EEG features were systematically explored to find the optimal bands for discriminating mild-AD patients from healthy normal elderly controls (Nold). As a result, three optimized bands (4-7 Hz, 8-15 Hz and 19-24 Hz) were identified, their individual classification performance was compared with the one obtained with their closet traditional band: theta, alpha and beta respectively, and, in the three cases, the use of the optimized bands led to a reduction of 20% or more on the classification error, when compared with the traditional bands. These results were later expanded in [24], where frequency band optimization was explored also for synchrony features, and for two different classification tasks, namely, Nold vs. mild cognitive impairment (MCI), and Nold versus AD. For both classification tasks the performance was improved 7% or more by the use of the non-traditional bands. Interestingly, the optimal bands were different for the two classifications tasks, showing that AD progression manifests itself across varying spectral regions. The results presented in [23] and [24] suggest that the optimal frequency bands for the study of AD may differ from traditional EEG bands, thus motivating the search for new features extracted from non-traditional frequency bands.

Motivated by outcomes obtained with the use of non-traditional frequency bands, in this paper we introduce a new type of feature for rsEEG-based AD assessment. The proposed features aim at characterizing the neuromodulatory deficit due to AD without being constrained by the use of traditional EEG bands. The features rely on a 2-dimensional representation called power modulation spectrogram and are computed based on the power available in three specific regions (or patches) of this 2-d representation. To gauge the advantages of the proposed features, performance comparison with several conventional rsEEG features is performed for three binary classification tasks, namely: Nold vs. AD, Nold vs. mild AD (early AD diagnosis task), and mild AD vs. moderate-to-severe AD (AD progression monitoring task). Next, a more difficult three-class task is explored: Nold vs. mild AD vs. moderate-to-severe AD. To validate the obtained findings, the biological origin of the proposed features was explored by using source localization and the results obtained were compared and corroborated with those reported in the literature of rsEEG-based AD assessment.

It is hoped that the methods presented herein can be used in clinical practice to assist clinicians with screening and/or pretesting of AD. Early detection of AD (i.e., the Nold vs. mild AD task), for example, is very important as it not only allows for measures to be taken early to improve the quality-of-life of patients, their families and caregivers, but may also maximize the benefits of administering AD-modifying drugs at early stages of the disease, thus increasing the odds of success. Lastly, monitoring progression of AD severity levels can open doors for lower-cost, remote AD monitoring tools.

The remainder of this paper is organized as follows: Section II describes the database employed for this work, the development of the proposed modulation-based features, the approached for feature selection and classification. Sections III and IV present the experimental results, and the discussion over theses. Lastly, conclusions are presented in Section V.

TABLE I
DEMOGRAPHIC CHARACTERISTICS OF PARTICIPANTS.

Group identifier	Subjects (female)	Age [years]	Education [years]	MMSE
Nold	20 (9)	68.0±8.6	10.1±5.5	28.5±1.7
AD1	19 (11)	74.1±5.5	5.6±2.8	19.4±5.3
AD2	15 (9)	75.0±11.8	4.1±3.8	12.8±5.0

p-values for age differences: Nold vs. AD1 p=0.012, Nold vs. AD2 p=0.061, AD1 vs. AD2 p=0.774.
p-values for education differences: Nold vs. AD1 p=0.003, Nold vs. AD2 p=0.0006, AD1 vs. AD2 p=0.207.

II. Materials and Methods

A. Participants

Fifty-four participants were recruited from the Behavioral and Cognitive Neurology Unit of the Department of Neurology and the Reference Center for Cognitive Disorders at the Hospital das Clinicas in São Paulo, Brazil. AD diagnosis and identification of progression level (mild-AD or moderate-to-severe AD), were performed by experienced neurologists according to NINCDS-ADRDA criteria [25] and classified based on the Brazilian version of the MMSE [26]. Participants were divided into three groups. The first group (Nold) consisted of 20 cognitively healthy normal elderly controls, the second group (AD1) of 19 mild-AD patients and the third group (AD2) of 15 patients with moderate-to-severe AD symptoms. Henceforth, the union of the AD1 and AD2 groups is referred to as AD group.

Inclusion criteria for the Nold group included a clinical dementia rating (CDR) score = 0 and a mini-mental state examination (MMSE) score $\geq$ 25, as well as no indication of functional cognitive decline. Inclusion criteria for the AD1 group, in turn, included $0.5 \leq$ CDR ≤ 1 and MMSE ≤ 24; lastly, inclusion criteria for the AD2 group were CDR score = 2 and MMSE ≤ 20. The p-values for age differences were: Nold vs. AD1 p=0.012, Nold vs. AD2 p=0.061, AD1 vs. AD2 p=0.774. The p-values for education differences were: Nold vs. AD1 p=0.003, Nold vs. AD2 p=0.0006, AD1 vs. AD2 p=0.207. Table I provides further details in the demographic characteristics of participants. For inclusion to the two AD groups, an additional criterion used was the presence of functional and cognitive decline over the previous 12 months based on detailed interviews with knowledgeable informants. Patients from the AD groups were also screened for diabetes mellitus, kidney disease, thyroid disease, alcoholism, liver disease, lung disease or vitamin B12 deficiency, as these can also cause cognitive decline. Ethics approval was obtained from the Research Ethics Office and participants consented to participate in the study.

B. EEG data acquisition and pre-processing

Twenty-channel EEG signals were acquired with the participants awake, relaxed, and with their eyes closed for at least eight minutes. EEG was recorded with 12-bit resolution and 200 Hz sampling frequency using Braintech 3.0

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

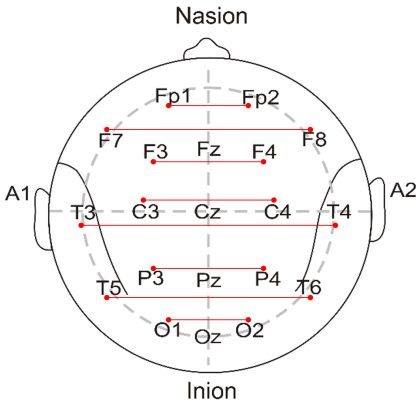

Fig. 1. EEG electrode location. Red lines indicate the electrodes pairs for the inter-hemispheric bipolar signals.

instrumentation (EMSA Equipamentos Médicos INC., Brazil). Scalp electrodes were placed according to the international 10-20 montage system; electrode impedance was kept below 10 kΩ, and bi-auricular (attached) electrodes were used as reference. EEG recordings were filtered with a zero-phase FIR bandpass filter with a bandwidth 0.5-45 Hz. Moreover, based on evidence of an inter-hemispheric disconnection with AD [11], [27], [13], [28], [29], we also explored the use of virtual inter-hemispheric bipolar signals which are the electric potential difference measured between a pair of electrodes symmetrically located in each hemisphere. The term "virtual" is used because these signals are mathematically computed as the subtraction of the two recorded unipolar signals involved [4]. Eight inter-hemispheric bipolar signals were computed: Fp1-Fp2, F7-F8, F3-F4, T3-T4, C3-C4, T5-T6, P3-P4 and O1-O2. Figure 1 presents the location of the electrodes, as well as the electrode pairs used in the computation of the inter-hemispheric bipolar signals.

C. EEG artifact handling

Recorded EEG signals were visually inspected by two experienced clinicians to identify several 8-second epochs free of eye blinking, drowsiness, muscle movements, or equipment-related artifacts; these artifact-free signals comprise the so-called "manually selected" EEG dataset. In addition to the "manually selected" dataset, the complete unprocessed EEG signals are also used, these signals are referred to as the "raw" EEG dataset. Lastly, following insights from [30], automated artifact removal was performed on the raw EEG dataset using the wavelet-enhanced independent components analysis (wICA) algorithm [31], [32], thus resulting in the "wICA" EEG dataset. This automated enhancement method showed to be the best for AD assessment [30].

With the wICA approach, the EEG signal is first decomposed into independent components (ICs), then the discrete wavelet transform is applied to each IC. The spectrotemporal representation of the IC is then thresholded in an attempt to recover only neural activity. Lastly, corrected ICs are projected to obtain the enhanced version of the EEG data. A complete description, as well as a comparative analysis between ICA and wICA is provided in [31]. The parameters used for wICA were 'cleaning artifact tolerance' = 1.25 and 'IC artifact detection threshold' = 4. Both parameters were empirically determined, and the degree of cleaning of the EEG signals was validated by visual inspection of wICA-enhanced EEG signals. The creation of the three EEG datasets (i.e., manually selected, raw and wICA) was performed using the EEGLAB [33] and wICA toolboxes [31].

D. EEG Modulation Spectrogram

In recent years, the modulation spectrogram has shown to be a relevant tool in the study of signals which present second-order periodicities as the result of AM processes [34], [13], [35], [36]. In our experiments, we used the power modulation spectrogram as a modulation-domain representation to characterize the periodicity in the power time series for different frequencies of the EEG signals. For each channel, $x(t)$, in the complete rsEEG recordings, the spectrotemporal representation, $X(t, f)$, was computed using the discrete-time continuous wavelet transform (heceforth referred to as CWT) with the complex Morlet wavelet as mother wavelet for frequencies from 1 to 45 Hz with 1-Hz step, and number of cycles $n_c = 6$. The complex Morlet was used as its time- and frequency-localization make it suitable for the analysis of rsEEG [37]. Then, the power spectrogram, ($|X(t, f)|^2$), was split into 8-second epochs (with 7 seconds overlap). In the case of the "manually selected" EEG datasets, only the epochs that were selected as artifact-free are considered for the following steps. For each epoch, its corresponding power modulation spectrogram was computed by applying the Fourier transform over the time dimension for each of the frequency components in the power spectrogram, as per:

$$X(f, f_{mod}) = \mathcal{F}_t \left\{ |X(t, f)|^2 \right\}. \quad (1)$$

Figure 2 presents the signal processing steps involved in the computation of the power modulation spectrogram. Note that if the two transformations involved in the computation of the power modulation spectrogram are invertible, processing can be done in the modulation spectral domain (e.g., filtering) and the time-domain signal can be reconstructed [35]. To allow for the comparison of the obtained power modulation spectrograms among epochs and subjects, each power modulation spectrogram was normalized by its total power, i.e., the sum of the square of the modulus of its elements. This process was performed for each of the (electrode and virtual) channels in the rsEEG recordings.

The comparison between two groups in the modulation domain was carried out as follows. For a given channel, the normalized power modulation spectrograms were aggregated first over all epochs within a subject, then over all subjects within each group, resulting in a collection of normalized power modulation spectrograms per group, as depicted in Figure 3a. Next, for each group a distribution of the power of each (f, f_{mod}) element in the normalized power modulation spectrograms is obtained as depicted in Figure 3b. From these two obtained distributions, the area under the receiver-operating-characteristic curve (AUC) [38], [39] was computed

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

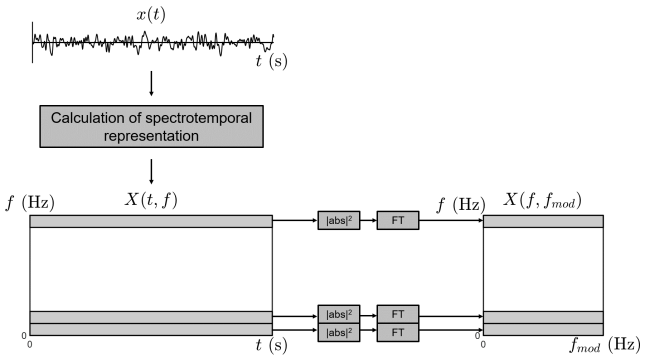

Fig. 2. Signal processing steps for the computation of the EEG power modulation spectrogram.

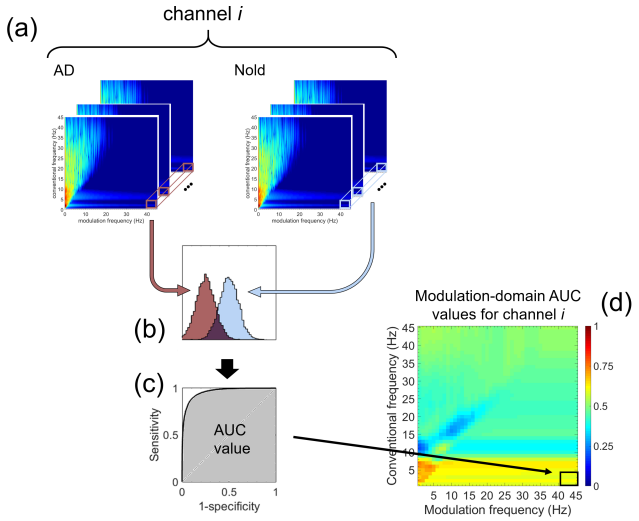

Fig. 3. Pipeline to compute the AUC values in modulation spectral domain.

as a metric of separability between groups, as shown in Figure 3c. Thus, for each (f, f_{mod}) element in the normalized power modulation spectrograms, an AUC value is obtained and represented in the modulation domain, as shown in Figure 3d.

Given the fact that the computation of the AUC values was performed for every channel in the rsEEG datasets, the result is a topographical plot, which presents the AUC values in the modulation domain for each channel for a specific group comparison task. Here, three group comparisons were performed: (*i*) Nold vs. AD, (*ii*) Nold vs. AD1, and (*iii*) AD1 vs. AD2. From the obtained AUC topographical plots (see Section III-A), we identified three specific regions (R_1, R_2 and R_3) in the modulation-domain representation that account for the main differences between the groups.

E. EEG feature extraction

To assess the discriminative value of the modulation representation for AD diagnosis, a new type of feature is proposed. More specifically, we measure the power of the three regions of interest in the power modulation spectrogram (more details to follow in Section III-A). The new "patch" features

TABLE II
LIST OF COMPUTED FEATURES FOR EACH 8-S EPOCH.

Feature group	**Features**
Modulation domain features (Computed for each of the 28 channels)	P_1
	P_2
	P_3
	P_1/P_3
	P_2/P_1
	P_2/P_3
Spectral features (Computed for each of the 28 channels)	*delta*
	theta
	alpha
	*alpha*1 (8 -10 Hz)
	*alpha*2 (10 - 12 Hz)
	beta
	delta-beta
	theta-beta
	gamma
Amplitude modulation rate-of-change (Computed for each of the 28 channels)	*delta-mdelta*
	theta-mdelta
	theta-mtheta
	alpha-mdelta
	alpha-mtheta
	beta-mdelta
	beta-mtheta
	beta-malpha
	beta-mbeta
	gamma-mdelta
	gamma-mtheta
	gamma-malpha
	gamma-mbeta
	gamma-mgamma

are defined as the integral of normalized power modulation spectrogram over the regions of interest, i.e.,:

$$P_i = \iint_{R_i} |X(f, f_{mod})|^2 df\, df_{mod}. \quad (2)$$

In addition to the three features, P_1, P_2 and P_3, derived from the three identified regions (R_1, R_2 and R_3, respectively), the ratios between these features, namely, $P_1/P_3, P_2/P_1$ and P_2/P_3 were also computed. These modulation-domain features were extracted for each of the normalized power modulation spectrograms obtained from 8-s epochs, for the 28 (electrode and virtual) channels. Eight second epochs have been show to be useful for modulation spectral analyses [13], [27].

Besides these newly proposed features, features commonly used in EEG-based AD assessment are also computed as benchmarks. More specifically, power spectral and AM rate-of-change features were used, as described in [14]. These conventional features were also computed over 8-s epochs for the 28 channels. Finally, the computed features were grouped into five feature sets: PSD (power spectral), MOD (amplitude modulation rate-of-change), RG3 (proposed features, power in the three regions of interests of the power modulation spectrogram and their ratios), and two combinations of these sets, namely, PSD+MOD and PSD+RG3. In our experiments we performed feature averaging over five consecutive epochs, as a way to improve the SNR of the extracted features, as suggested in [27]. Table II presents the list of the computed features for the manually selected, raw, and wICA EEG datasets.

F. Feature selection and classification

Feature selection was used as high-dimensionality feature vectors often lead to bias and overfitting in classification when

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

limited data is available. By setting aside a portion of the available data for feature selection, we assure greater strictness of the test and minimize any "leakage" of insights from the test set into the training set. Here, 25% of the available data per participant was set aside for feature engineering and feature selection, as suggested in [30]. A simple analysis of variance (ANOVA) between feature value and group label was performed for feature selection and the top-24 features based on the ANOVA F-values were chosen. Only top-ranked 24 features were used to remain inline with previous classifiers reported in the literature [27].

For classification, a support vector machine (SVM) classifier was used with the following parameters: radial basis function (RBF) kernel, $gamma = 1/number\_of\_features$, and a default regularization coefficient value of $C = 1$. The open-source scikit-learn (machine learning for Python) toolbox [40] was used for feature selection and classification. Classifier performance was evaluated using the leave-one-subject-out (LOSO) approach for cross-validation (CV). In LOSO CV, from a dataset comprised of data from N participants, the data from $N - 1$ subjects is used to train the classifier while data from the remaining subject is used for testing. This process is repeated for each participant in the dataset. The LOSO CV approach was further extended by bootstrapping. Under this LOSO+Bootstrapping CV setup, for each subject, the classifier is trained with data from $N - 1$ subjects that is randomly sampled with repetition 10 times.

Finally, the label for each subject is defined based on majority vote, i.e., if 50% or more of the subject's epochs belong to a given class the subject is assigned to that class. Four classification tasks were performed: (*i*) Nold vs. AD, (*ii*) Nold vs. AD1, (*iii*) AD1 vs. AD2, and (*iv*) Nold vs. AD1 vs. AD2. As figures of merit, accuracy and F1-score were computed. Moreover, as the groups were not age-matched and age has been shown to be a confounding factor, we further assess the impact of age on the proposed models. To this end, we include age into the feature selection and classification stages and compare performance to the case without.

G. Source localization

In order to better understand the biological origin of the differences observed in modulation spectral power between the different groups, cortical source localization was performed. To this end, the manually selected EEG signals were filtered in the modulation domain to preserve one of the three modulation regions (i.e., R_1, R_2, R_3), as well as all three combined ($R_1 + R_2 + R_3$). Filtering in the modulation domain was performed as described in [35]. The modulation-domain filtered rsEEG recordings were then used to estimate the electric neuronal activity distribution with the eLORETA method [41] using the LORETA-KEY software provided by The KEY Institute of Brain-Mind Research, University Hospital of Psychiatry, Switzerland. The head model used and the electrode coordinates are based on the Montreal Neurological Institute (MNI) average MRI brain map, MNI152 template [42]. The solution space was limited to the cortical gray matter, including 6239 voxels with spatial resolution of 5 mm^3. The corresponding Brodmann areas (BAs) are reported using the MNI space with correction to Talairach space [43].

The eLORETA cortical current source density solution was computed for eight frequency bands: delta (1.5-6 Hz), theta (6.5-8 Hz), alpha1 (8.5-10 Hz), alpha2 (10.5-12 Hz), beta1 (12.5-18 Hz), beta2 (18.5-21 Hz), beta3 (18.5-30 Hz) and full-band (1.5-30 Hz). It is important to emphasize that eLORETA was used with traditional EEG frequency bands in order to directly compare the obtained results with the ones reported in the AD literature. The statistical analysis of the localization solutions was performed with the statistical non-parametric mapping (SnPM) method. The difference in cortical source localization between Nold and AD groups was assessed for each of the eight frequency bands with voxel-by-voxel independent F-ratio tests, based on eLORETA log-transformed current source density power. Cortical voxels with significant differences were identified by means of a non-parametric randomization procedure (5000 randomizations) in the three-dimensional statistical mapping. The mean source power in each voxel and the distribution in the permutated values was compared, with a threshold set at a p=0.05 significance level after correction for multiple comparisons across all voxels and frequencies [44]. It is important to emphasize that while eLORETA has commonly been applied with $64+$ EEG electrode setups [45], reliable results have been reported in the literature with as few as 16 electrodes [46], with the majority of the AD studies relying on 19 electrodes [8].

III. Experimental Results

A. Comparison between groups

From each comparison of the normalized power modulation spectrograms between the Nold and AD groups an AUC topographic plot was obtained, as described in Section II-D. Figure 4 depicts the AUC topographic plot for 25% subset of the manually selected EEG dataset; the AUC topographic plots for the raw and wICA datasets are available as supplementary material. In Figure 4, it can be observed that there are regions in the modulation domain where the AUC values are far from 0.5. More specifically, there are three regions (R_1, R_2 and R_3) that stand out. These regions are depicted by Figure 5a, which corresponds to the average over all 20 electrodes shown in Figure 4 which was obtained for the (subset) manually selected EEG data. In a similar manner, Figures 5b and 5c present the average AUC values across all 20 electrodes for the (subset) raw and wICA-enhanced datasets, respectively. The averaging across electrodes provides an overall view of the differences between groups in the modulation domain, thus facilitating the identification of the regions with higher AUC values. However, this procedure removes spacial information, thus further motivating the source localization step used next.

The results presented in Figures 4 to 5 were obtained from the comparison between the groups Nold and AD, the latter comprised by individuals from the combined AD1 and AD2 groups. Further exploration of the different AD groups was also performed. Figure 6 depicts the average across electrodes obtained from the comparisons for Nold vs. AD1, Nold vs. AD2, and AD1 vs. AD2 for the manually selected EEG dataset.

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

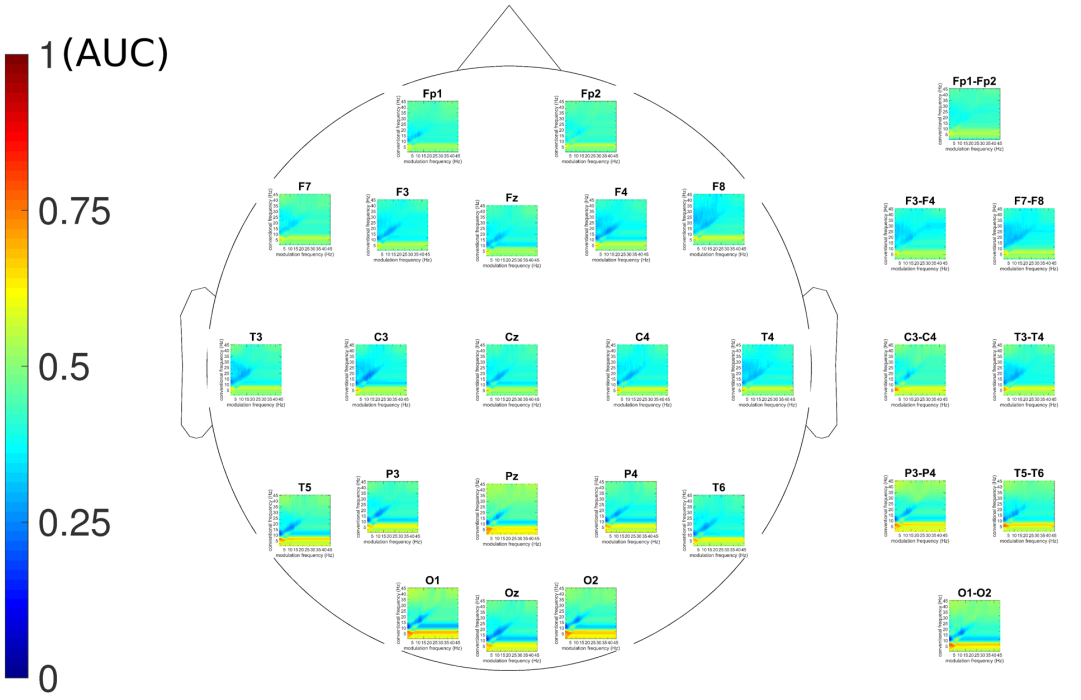

Fig. 4. AUC topographic plot for Nold vs. AD, derived from the manually selected EEG dataset.

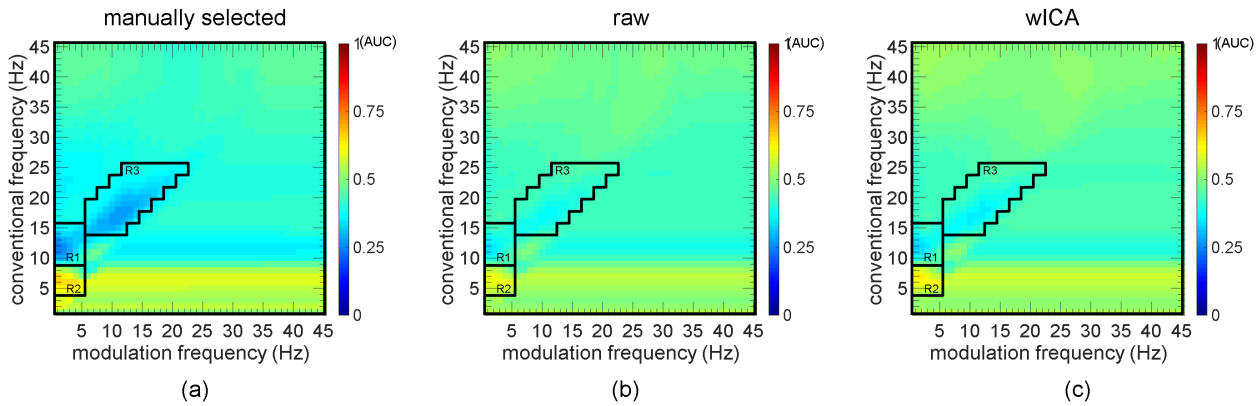

Fig. 5. AUC values averaged across monopolar electrodes from Nold vs AD comparison with: (a) manually selected, (b) raw, and (c) wICA-enhanced EEG recordings.

B. Classification performance

EEG features were computed for each of the EEG datasets, as described in Section II-E. With these features, three binary and one three-class classification tasks were carried out: (*i*) Nold vs. AD, (*ii*) Nold vs. AD1, (*iii*) AD1 vs. AD2, and (*iv*) Nold vs AD1 vs AD2. Tables III to VI report the mean and standard deviation of the accuracy and F1-score achieved in each classification tasks for each combination of EEG datasets and feature sets (not including age as a feature). In these Tables, an asterisk indicates the cases where the inclusion of age in the feature set had a statistically significant improvement ($p<0.05$) in the figures of merit. For the manually selected EEG dataset, the RG3 feature set was the best performer for three of the four classification task, and the second best for the remaining classification task (Nold vs. AD1). Moreover, these performances were not affected by the introduction of age as feature. Table VII shows the top-24 selected features the RG3 feature set (without age as a feature) in each classification task for the manually selected dataset. Feature names are reported according the nomenclature presented in Table II.

C. Source localization

As mentioned previously, EEG signals were filtered in the modulation domain with the goal of preserving only the modulation regions of interest. Using the rsEEG recordings from the Nold and AD groups, source localization was per-

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

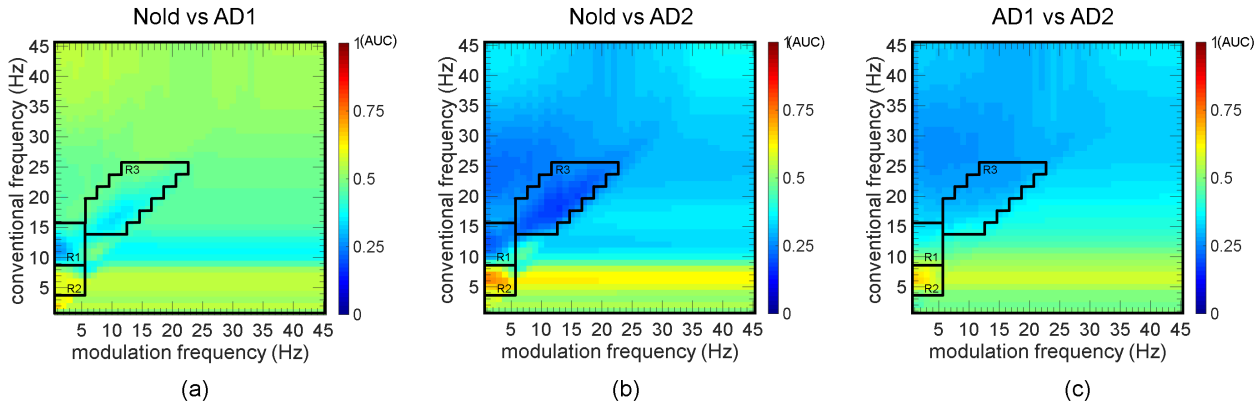

Fig. 6. AUC values averaged across monopolar electrodes in the manually selected EEG dataset for (a) Nold vs. AD1, (b) Nold vs. AD2, and (c) AD1 vs. AD2 comparisons.

TABLE III

MEAN AND STANDARD DEVIATION (SHOWN BETWEEN PARENTHESES) FOR ACCURACY AND F1-SCORE FOR THE NOLD VS. AD CLASSIFICATION TASK. VALUES IN BOLD INDICATE THE BEST PERFORMANCE FOR EACH EEG DATASET. THE * INDICATE THE CASES WHERE THE INCLUSION OF AGE IN THE FEATURE SET HAD A STATISTICALLY SIGNIFICANT ($p<0.05$) IMPROVEMENT IN THE FIGURES OF MERIT.

	Feature sets									
	PSD		MOD		RG3		PSD+MOD		PSD+RG3	
EEG datasets	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
manually sel.	84.4 (1.2)	82.0 (1.4)	74.1 (3.4)	70.8 (3.8)	**88.1 (1.9)**	**86.2 (2.3)**	84.8 (0.7)	82.5 (0.7)	79.1 (1.9)	75.6 (1.8)
raw	73.1 (2.5)	70.3 (2.7)	63.1* (1.7)	53.4* (1.9)	**78.1 (1.4)**	**73.2 (1.9)**	72.8* (1.2)	67.4* (1.5)	76.7 (1.2)	72.2 (1.0)
wICA	70.6* (1.0)	67.1* (1.2)	60.2* (3.0)	50.1* (2.5)	72.2 (2.3)	67.7 (2.7)	73.5 (1.9)	69.1 (2.7)	**76.3 (0.7)**	**72.2 (1.0)**

TABLE IV

MEAN AND STANDARD DEVIATION (SHOWN BETWEEN PARENTHESES) FOR ACCURACY AND F1-SCORE FOR THE NOLD VS. AD1 CLASSIFICATION TASK. VALUES IN BOLD INDICATE THE BEST PERFORMANCE FOR EACH EEG DATASET. THE * INDICATE THE CASES WHERE THE INCLUSION OF AGE IN THE FEATURE SET HAD A STATISTICALLY SIGNIFICANT ($p<0.05$) IMPROVEMENT IN THE FIGURES OF MERIT.

	Feature sets									
	PSD		MOD		RG3		PSD+MOD		PSD+RG3	
EEG datasets	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
manually sel.	72.6 (2.6)	72.5 (2.6)	**78.7*** (**2.3**)	**78.7*** (**2.3**)	72.6 (1.6)	72.5 (1.7)	72.1 (2.9)	72.0 (2.9)	69.0 (1.4)	68.9 (1.4)
raw	**73.3 (2.4)**	**73.3 (2.3)**	45.6* (5.0)	45.3* (5.2)	62.1* (2.5)	61.9* (2.6)	67.9* (1.3)	67.9* (1.3)	65.4* (1.7)	65.4* (1.7)
wICA	66.4* (2.1)	66.4* (2.1)	47.2* (2.1)	46.8* (2.1)	63.6* (1.9)	63.5* (2.0)	**71.0 (2.8)**	**71.0 (2.8)**	66.7 (3.8)	66.6 (3.8)

TABLE V

MEAN AND STANDARD DEVIATION (SHOWN BETWEEN PARENTHESES) FOR ACCURACY AND F1-SCORE FOR THE AD1 VS. AD2 CLASSIFICATION TASK. VALUES IN BOLD INDICATE THE BEST PERFORMANCE FOR EACH EEG DATASET.

	Feature sets									
	PSD		MOD		RG3		PSD+MOD		PSD+RG3	
EEG datasets	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
manually sel.	67.6 (1.9)	67.3 (1.9)	64.7 (2.9)	64.2 (2.9)	**75.6 (2.3)**	**74.9 (2.5)**	63.8 (3.0)	62.8 (2.9)	73.8 (0.9)	73.2 (1.0)
raw	70.6 (2.9)	69.8 (3.1)	71.8 (2.0)	70.2 (2.3)	**78.8 (1.2)**	**74.9 (2.5)**	69.7 (3.0)	68.2 (1.9)	74.7 (1.4)	73.7 (1.6)
wICA	68.5 (2.3)	67.7 (2.3)	68.5 (1.9)	67.9 (2.0)	**76.8 (2.1)**	**76.1 (2.1)**	72.9 (2.2)	72.4 (2.1)	70.6 (0.0)	66.6 (0.0)

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

TABLE VI
MEAN AND STANDARD DEVIATION (SHOWN BETWEEN PARENTHESES) FOR ACCURACY AND F1-SCORE FOR THE NOLD VS. AD1 VS. AD2 CLASSIFICATION TASK. VALUES IN BOLD INDICATE THE BEST PERFORMANCE FOR EACH EEG DATASET.

	Feature sets									
	PSD		MOD		RG3		PSD+MOD		PSD+RG3	
EEG datasets	Acc	F1	Acc	F1	Acc	F1	Acc	F1	Acc	F1
manually sel.	63.0 (1.7)	63.0 (1.8)	47.4 (2.4)	48.1 (2.3)	**69.3** (**1.5**)	**68.2** (**1.5**)	58.0 (2.2)	57.5 (2.3)	64.4 (2.0)	63.0 (2.0)
raw	**58.5** (2.2)	**58.4** (2.1)	35.6 (1.8)	36.7 (1.5)	50.2 (1.3)	50.3 (1.4)	53.3 (1.6)	53.9 (1.7)	55.7 (1.0)	55.3 (1.0)
wICA	51.5 (2.0)	51.8 (2.0)	38.3 (2.0)	41.0 (2.0)	50.6 (1.4)	50.7 (1.4)	52.2 (1.4)	53.4 (1.2)	**55.6** (**2.2**)	**56.1** (**2.3**)

TABLE VII
TOP-24 SELECTED FEATURES FOR EACH CLASSIFICATION TASK WITH THE MANUALLY-SELECTED EEG DATASET AND RG3 FEATURE SET. THE LAST SECTIONS OF THE TABLE SHOW THE NUMBER OF FEATURES THAT BELONG TO EACH CATEGORY.

	Best classification performance per classification task			
	Nold vs AD	Nold vs AD1	AD1 vs AD2	Nold vs AD1 vs AD2
Ranking	RG3	RG3	RG3	RG3
1	P_2/P_1-O1	P_2/P_1-O1-O2	P_3-T3-T4	P_2/P_1-O1
2	P_2/P_1-O1-O2	P_2/P_1-O1	P_2/P_3-T3	P_2/P_1-T5
3	P_2/P_1-T5	P_2/P_1-O2	P_1/P_3-C3	P_2/P_3-O2
4	P_2/P_1-O2	P_2/P_1-T5	P_2/P_3-C3	P_2/P_3-T5
5	P_2/P_1-T5-T6	P_2-P3-P4	P_1/P_3-Cz	P_2/P_3-O1
6	P_2/P_1-Oz	P_2/P_1-P3-P4	P_1/P_3-F4	P_3-T3-T4
7	P_2/P_1-P3-P4	P_2/P_1-P4	P_2/P_3-T5	P_2/P_1-O1-O2
8	P_2-P3-P4	P_2/P_1-T5-T6	P_2/P_3-T3-T4	P_2/P_1-O2
9	P_2/P_1-Pz	P_1-O1	P_2/P_3-Oz	P_2/P_3-F4
10	P_2/P_1-P3	P_2/P_1-Oz	P_2/P_3-Cz	P_2/P_1-T5-T6
11	P_2/P_1-P4	P_1-O1-O2	P_1/P_3-C4	P_2/P_3-T3
12	P_2/P_3-O2	P_1-Oz	P_2/P_3-O1	P_2/P_1-Pz
13	P_2-T5	P_2/P_1-C3-C4	P_2/P_3-O2	P_2/P_3-Oz
14	P_2/P_3-O1	P_2/P_1-P3	P_2/P_3-F4	P_2/P_1-P3
15	P_2/P_3-O1-O2	P_2/P_1-C3	P_3-T3	P_2/P_3-T3-T4
16	P_2/P_1-C3	P_2/P_1-F4	P_3-T5	P_2/P_3-C3
17	P_2/P_3-T5	P_1-O2	P_2/P_3-F3	P_2/P_3-F7-F8
18	P_2/P_3-T5-T6	P_2/P_1-Pz	P_1/P_3-T5	P_2/P_3-T5-T6
19	P_2/P_3-F7-F8	P_2/P_3-F7-F8	P_1/P_3-C3-C4	P_2/P_1-Oz
20	P_2/P_1-T6	P_2/P_3-O1-O2	P_2/P_3-Pz	P_2/P_1-P3-P4
21	P_2-O2	P_2-T5	P_1-C3-C4	P_2-T5
22	P_2/P_1-C3-C4	P_2-O2	P_2/P_3-Fz	P_2/P_1-P4
23	P_2/P_1-C3	P_2-O1	P_2/P_1-T3	P_2/P_3-O1-O2
24	P_1-Pz	P_2/P_1-T6	P_2/P_3-P3	P_2/P_3-P3
Number of features per patch				
P_1	1	4	1	0
P_2	3	4	0	1
P_3	0	0	3	1
P_1/P_3	0	0	6	0
P_2/P_1	14	14	1	10
P_2/P_3	6	2	13	12
Number of features per brain region				
Frontal	1	2	4	2
Central	3	2	7	1
Temporal	6	4	8	8
Parietal	6	5	2	5
Occipital	8	11	3	8
Number of features from inter-hemispheric channels				
	8	8	4	8

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

TABLE VIII
BRODMANN AREAS WITH STATISTICALLY SIGNIFICANT DIFFERENCE (p<0.05) BETWEEN NOLD AND AD COHORTS. THE ASTERISK INDICATES p<0.01. LL = LIMBIC LOBE, OC = OCCIPITAL LOBE, PL = PARIETAL LOBE, TL = TEMPORAL LOBE, AND FL = FRONTAL LOBE.

Region preserved	***Band*** **Lobe (BAs)**
R1	*high-alpha (AD < Nold)* LL (18,19,23,29,30,31,37) OL (7,17*,18*,19*,23,30,31,37) PL (7,19,31,39,40) TL (13,19,20,21,22,37,39,40)
R2	*theta (AD > Nold)* FL (5,6,31,43,44) LL (20,23,27,28,29,30,31*,34,35,36) OL (19,23,31) PL (5,7,19,31,39,40) TL (13,20,21,22,38,39,41)
R3	*high-alpha (AD < Nold)* LL (18,19,23,29,30,31) OL (7,17,18,19,23,30,31,37) PL (19,31) TL (19,20,37,39)
R1+R2+R3	*high-alpha (AD < Nold)* LL (18,19,23,29,30,31,37) OL (7,17*,18*,19*,23*,30*,31,37*) PL (7,19*,31,39,40) TL (13,19,20,21,22,37,39,40)

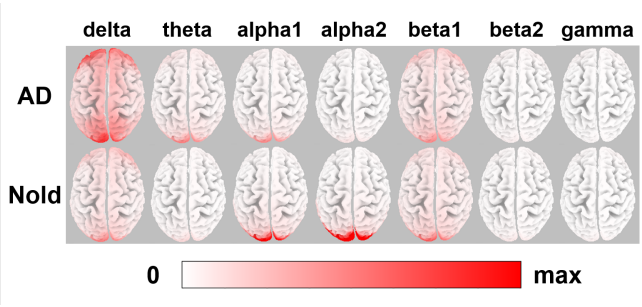

Fig. 7. Group average of normalized LORETA solutions, scaled to the max value (average of alpha2 Occipital sources in Nold cohort)

formed with eLORETA for each individual region alone, as well as all three regions combined. The BAs that presented statistically significant differences (p<0.05) are listed in Table VIII. Additionally, the LORETA solutions obtained for each conventional frequency band for the Nold and AD cohorts are depicted in Figure 7. These source localization solutions were obtained with the rsEEG signals filtered in the modulation domain to keep R1+R2+R3.

IV. DISCUSSION

A. Classification performance

As mentioned previously, direct comparison with other studies reported in the literature is difficult, as there is no standard method of pre-processing, feature selection or testing setup for the task at hand. In this work, for example, LOSO CV was used to avoid the optimistic bias introduced by k-fold CV (e.g., see [14]). In particular, LOSO combined with bootstrapping was used. In Tables III-V, it can be seen that for most of the classification tasks, the proposed RG3 features resulted in the highest performance. Given the fact that not all the cohorts were age-matched, the classification tasks were also evaluated by using age as a feature in each dataset. As a result, the classification performance for tasks Nold vs. AD and Nold vs. AD1 were improved by the use of age as a feature. In the case of Nold vs. AD, for example (Table III), the addition of age accounted for an average improvement of 6.1% in accuracy and 8.2% in F1-score in four cases. Notwithstanding, despite the gains seen by adding age as a feature, the obtained accuracy still remained below that achieved by the best performing classifier that did not rely on age as a feature.

For the classification task Nold vs. AD1, on the other hand (Table IV), inclusion of age as a feature showed to improve performance for eight of the fifteen cases tested. An average improvement of 8.6% in accuracy and 8.6% in F1-score was seen and the highest achieving classifier (MOD features and manually selected epochs) relied on age as a feature. Manually selected epochs, however, require human intervention, thus are not automated. In the case of automated methods, however, the best performing classifiers did not rely on age as a feature (e.g., PSD+MOD feature and wICA enhancement).

Lastly, for the classification task AD1 vs. AD2, and Nold vs. AD1 vs. AD2 (Tables V and VI, respectively), the addition of age as feature did not modify classification performance, thus suggesting that age did not serve as a bias for classification in these scenarios.

In general, the proposed RG3 feature set outperformed the conventional MOD features across most classification tasks, thus showing the importance of non-conventional bands, both in the frequency and modulation frequency domains, for AD diagnostics and severity level monitoring. Interestingly, when looking at the RG3 feature set alone, wICA processing resulted in equal or lower performance relative to using the raw EEG signals; this was true for all four examined classification tasks. Such findings suggest that wICA processing may remove discriminatory information that is captured by the spectrotemporal patches used in the RG3 features, despite this not being visually observed from the AUC plots of Fig.5.

Moreover, from the three different 2-class problems evaluated, the proposed RG3 feature achieved lower performance relative to the benchmarks only for the Nold vs. AD1 task, thus suggesting that further developments (e.g., alternate “patches”) are still needed for very early detection of AD. In fact, from the plot in Figure 6a, such results were expected as the AUC values for this classification task were lower than in the other tasks. Notwithstanding, while the MOD feature set achieved the highest accuracy in this task, it relied on manually selected epochs and performance quickly dropped once raw or wICA signals were used, thus suggesting their sensitivity to EEG artifacts. Lastly, for the more difficult three-class problem, the proposed RG3 set alone resulted in the highest accuracy and F1 score, with values close to 70% with manually selected epochs. While accuracy quickly dropped for all tested feature sets in the automated settings, the proposed set showed to complementarity to the benchmark PSD features and resulted in further improvements for the wICA case.

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.2019.2953475, IEEE Journal of Biomedical and Health Informatics

B. Source localization

By analyzing the modulation domain regions depicted in Figure 5, it can be noted that filtering the signal in the modulation domain for regions R_1 or R_3 conserves the information that roughly corresponds to the alpha2 and beta1 bands, and the region R_2 conserves information related to the theta band. As can be seen from Table VIII, the reported BAs are accordant with BAs reported in the AD literature, in particular for the Nold vs AD task where LORETA methods have been widely used (e.g., [44], [47], [48], [49], [50], [51]).

For example, the relationship between the degree of atrophy of the brain structures, specifically in gray matter (measured with MRI) and the power of the cortical solutions (obtained by using LORETA on rsEEG recordings) was explored in [52] for different stages of AD. In that study, it was found that the atrophy of the cortical gray matter had a negative correlation with the power of cortical delta sources in the central, frontal, parietal, occipital, temporal and limbic regions. A positive correlation was also reported between the atrophy of the cortical gray matter and power of cortical alpha1 sources in the parietal, occipital and limbic regions. Moreover, the progression of the hippocampus atrophy was shown to be correlated with decreased power of cortical alpha sources (estimated by using LORETA), in the course of AD [53]. As such, the results presented in this work suggest that the differences observed in the modulation domain are related to the characteristic neurodegeneration of AD. Given these insights, it is expected that the analysis of these regions in the modulation domain can improve rsEEG-based AD diagnosis. Moreover, filtering the EEG signals in the modulation domain has the potential of enhancing the rsEEG signal as only components of interest for AD diagnosis are kept, while rsEEG components not related to AD or that may be corrupted by artifacts are removed. Using this technique for EEG enhancement is part of future work.

C. Feature importance

An in-depth analysis of the top-24 features presented in Table VII shows that from the six different patch features proposed in this work, P_2/P_1 and P_2/P_3 accounted for 75% of the features. However, their importance was not the same for each of the classification tasks. For example, P_2/P_1 showed to be the most relevant for the Nold vs. AD, and Nold vs. AD1 tasks. In turn, it just appeared once in the top features for the AD1 vs. AD2 classification task, where P_2/P_3 was the most discriminative feature. In the case of the three-class task, both P_2/P_1 and P_2/P_3 features amounted to roughly half the top features each. These changes across tasks was expected, as the different patches conveyed dfferent discriminatory power based on severity level, as shown in Fig. 6.

Regarding the regions of the brain related to the top features, for classification tasks that comprehend both AD cohorts, i.e. Nold vs. AD, and Nold vs. AD1 vs. AD2, most of the features were derived from electrodes over the temporal, parietal and occipital regions, thus in agreement with the analysis of atrophy in cortical surfaces during AD, which indicates that AD patients generally show changes in temporal and parietal regions. [54]. On the other hand, classification tasks involving different stages of AD, i.e. N vs AD1, and AD1 vs. AD2, the most important brain regions (according to the number of features related to them) where parietal and occipital for N vs. AD1, and temporal and central for AD1 vs. AD2. This is in accordance with the cortical signature of the progression of AD, in which regions such as the medial temporal, inferior temporal, temporal pole, inferior parietal, and posterior cingulate/precuneus are typically affected in early stages of AD. With the progression of AD, there is an expansion in the atrophy from these regions towards regions such as the superior parietal and frontal cortex [55].

Lastly, as can be seen from the Table, for all tasks except AD1 vs AD2, one third of the top-24 features were derived from virtual inter-hemispheric bipolar signals. This inter-hemispheric disconnect has been well documented in AD [11]. Interestingly, for the disease severity detection task, only four of the top-24 features were derived from the bipolar signals. Previous studies have reported an inter-hemispheric disconnection problem already at the mild cognitive impairment stage [56]. As such, inter-hemispheric bipolar signals may be useful for detection tasks that seek to discriminate Nold from AD patients; however, they may not be as effective to discriminate between disease severity levels.

D. Towards low-density devices for AD monitoring

For AD research and clinical applications, EEG devices typically have 16 or more electrodes (often 64+), thus they are not only hard to transport and expensive to fund, but also have long preparation times which might evoke undesired cognitive states such as drowsiness, fatigue, frustration or stress in elderly participants. As a consequence, low-density EEG devices may be desirable, particularly if portable and remote applications are sought [8]. A notable characteristic of the AUC topographic plot presented in Figure 4 is that the differences between groups are not uniform across the scalp, where the electrodes over the occipital, parietal and temporal lobes present the best AUC values. This behaviour is also seen in Table VII, where around 80% of the features were derived from electrodes over the occipital, parietal and temporal lobes. Such findings suggest that low-density (potentially portable) EEG devices could be designed specifically for AD diagnosis and progression monitoring based on rsEEGs.

E. Study limitations and future work

Small sample sizes has been a reported limitation of AD databases [8]. Here, data from 54 subjects was available, thus methods of splitting the data can result in large performance variability. To minimize this impact, when selecting the 25% withheld set (for feature selection and feature engineering), sampling was done from random places in the rsEEG recording and this partitioning was done ten times. We found that accuracy differences were not significant between partitions. To further minimize any effects of sample size, bootstrapping was also used within our LOSO CV setup.

Moreover, since a small amount of data (25%) was set aside for feature engineering, selection of the RG3 regions was performed visually in an ad-hoc manner based on AUC values

obtained between Nold and AD groups. This may explain the lower performance achieved in the Nold vs AD1 case. To overcome this limitation, data-driven region selection could be performed for each task separately. Such methods, however, typically rely on large amounts of data.

Motivated be previous works, the proposed features have been computed from 8-second epochs [13], [27]. While this duration was shown optimal for the MOD features, alternate size windows may be better suited for the proposed RG3 features. Similarly, the Morlet wavelet was used here as it has shown to provide time- and frequency-localization properties useful for rsEEGs [37]. Notwithstanding, alternate wavelet families may be better suited for the task at hand. An in-depth analysis of the effect of window size and/or wavelet mother on AD monitoring performance is left for a future study.

Brain atrophy has also shown to be a marker that is correlated with presence and progression of AD [57]. As such, diverse works [52], [53], [58], [59] have explored the relationship between traditional EEG features and changes in the brain volumes. Future exploration of the proposed features and brain atrophy could also lead to interesting insights.

Lastly, age and education have been shown to be important co-factors in AD. A systematic review on the relationship between education and AD [60] found that 58% of the reviewed studies reported a significant effect of lower education on the incidence of AD. While the most supported explanation of this is the cognitive reserve hypothesis [61], life experiences and/or biological factors can also play a role [62]. Age, in turn, is known to be a stronger co-factor. Indeed, when using age as a feature, it was found in one scenario (Nold vs. AD1) that it contributed significantly towards classification accuracy. Future works should take age and education into account, be it by the data collection protocol, the feature/classifier development, or in the significance analyses.

V. Conclusion

In this paper, we propose a new type of rsEEG feature for the diagnostics and progression monitoring of Alzheimer's disease. The features are derived from the modulation signal representation with the goal of characterizing the neuromodulatory deficit of the EEG signals due to AD. Unlike existing features, the proposed ones are not constrained by the use of traditional EEG bands. In particular, the new features are shown to accurately monitor disease severity progression, and to perform better with raw EEG signals than with wICA-processed signals. Moreover, they are shown to outperform conventional features across four classification tasks, namely: Nold vs. AD, AD1 vs. AD2, and Nold vs. AD1 vs. AD2. Lastly, we explored the relationship between the proposed features and their biological origin by estimating the cortical sources related to the proposed features. This exploration showed that the associated brain regions correspond to areas that have been previously linked to AD. While the proposed features were developed within the context of AD dianosis and severity level assessment, they could provide new insights for other allied applications based on EEG, such as neurofeedback and brain-computer interfaces.

Acknowledgment

The authors would like to thank Prof. Francisco Fraga (UFABC) and Dr. Renato Anghinah (USP) for making the EEG database available and for earlier collaborations on EEG signal processing, as well as NSERC for financial support for this work.

References

[1] M. Prince, A. Wimo, M. Guerchet, G.-C. Ali, Y.-T. Wu, and M. Prina, "World Alzheimer Report 2015, The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends," *Alzheimer's Disease International*, p. 87, 2015.

[2] J. Cummings and N. Fox, "Defining Disease Modifying Therapy for Alzheimer's Disease," 2017.

[3] C. R. Jack, M. S. Albert *et al.*, "Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, vol. 7, no. 3, pp. 257–262, May 2011.

[4] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford; New York: Oxford University Press, 2006.

[5] F. Vecchio, C. Babiloni *et al.*, "Resting state cortical EEG rhythms in Alzheimer's disease: Toward EEG markers for clinical applications: A review," *Supplements to Clinical Neurophysiology*, vol. 62, pp. 223–236, 2013.

[6] R. Lizio, F. Vecchio, G. B. Frisoni, R. Ferri, G. Rodriguez, and C. Babiloni, "Electroencephalographic Rhythms in Alzheimer's Disease," *International Journal of Alzheimer's Disease*, vol. 2011, pp. 1–11, 2011.

[7] What Are the Signs of Alzheimer's Disease? http://www.nia.nih.gov/health/what-are-signs-alzheimers-disease.

[8] R. Cassani, M. Estarellas, R. San-Martin, F. J. Fraga, and T. H. Falk, "Systematic Review on Resting-State EEG for Alzheimer's Disease Diagnosis and Progression Assessment," *Disease Markers*, vol. 2018, 2018.

[9] A. W. Laxton, D. F. Tang-Wai *et al.*, "A phase I trial of deep brain stimulation of memory circuits in Alzheimer's disease," *Annals of Neurology*, vol. 68, no. 4, pp. 521–534, Oct. 2010.

[10] N. Hansen, "Brain Stimulation for Combating Alzheimer's Disease," *Frontiers in Neurology*, vol. 5, Jun. 2014.

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

[12] T. H. Ferreira-Vieira, I. M. Guimaraes, F. R. Silva, and F. M. Ribeiro, "Alzheimer's disease: Targeting the cholinergic system," *Current neuropharmacology*, vol. 14, no. 1, pp. 101–115, 2016.

[13] T. H. Falk, F. J. Fraga, L. Trambaiolli, and R. Anghinah, "EEG amplitude modulation analysis for semi-automated diagnosis of Alzheimer's disease," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1–9, 2012.

[14] R. Cassani, T. H. Falk, F. J. Fraga, M. Cecchi, D. K. Moore, and R. Anghinah, "Towards automated electroencephalography-based Alzheimer's disease diagnosis using portable low-density devices," *Biomedical Signal Processing and Control*, vol. 33, pp. 261–271, Mar. 2017.

[15] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, pp. 2266–2301, Oct. 2005.

[16] J. Dauwels, K. Srinivasan *et al.*, "Slowing and Loss of Complexity in Alzheimer's EEG: Two Sides of the Same Coin?" *International Journal of Alzheimer's Disease*, vol. 2011, pp. 1–10, 2011.

[17] H. Garn, M. Waser *et al.*, "Quantitative EEG in Alzheimer's disease: Cognitive state, resting state and association with disease severity," *International Journal of Psychophysiology*, vol. 93, no. 3, pp. 390–397, Sep. 2014, 00000.

[18] D. Moretti, G. Frisoni *et al.*, "MCI patients' EEGs show group differences between those who progress and those who do not progress to AD," *Neurobiology of Aging*, vol. 32, no. 4, pp. 563–571, Apr. 2011.

[19] A. Alberdi, A. Aztiria, and A. Basarab, "On the early diagnosis of Alzheimer's Disease from multimodal signals: A survey," *Artificial Intelligence in Medicine*, vol. 71, pp. 1–29, Jul. 2016.

[20] J. Dauwels, F. Vialatte, and A. Cichocki, "Diagnosis of Alzheimer's Disease from EEG Signals: Where Are We Standing?" *Current Alzheimer Research*, vol. 7, no. 6, pp. 487–505, Sep. 2010.

[21] D. Hoyer, R. Bauer *et al.*, "Specific monitoring of neonatal brain function with optimized frequency bands," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 5, pp. 40–46, 2001.

[22] Z. Dvey-Aharon, N. Fogelson, A. Peled, and N. Intrator, "Schizophrenia detection and classification by advanced analysis of eeg recordings using a single electrode approach," *PloS one*, vol. 10, no. 4, p. e0123033, 2015.

[23] M. Elgendi, F. Vialatte, A. Cichocki, C. Latchoumane, Jaesung Jeong, and J. Dauwels, "Optimization of EEG frequency bands for improved diagnosis of Alzheimer disease," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, Aug. 2011, pp. 6087–6091.

[24] 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.

[25] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. 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, 1984.

[26] S. Brucki, R. Nitrini, P. Caramelli, P. H. Bertolucci, and I. H. Okamoto, "Suggestions for utilization of the mini-mental state examination in Brazil," *Arquivos de neuro-psiquiatria*, vol. 61, no. 3B, pp. 777–781, 2003.

[27] F. J. Fraga, T. H. Falk *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 (ICASSP)*, May 2013, pp. 1207–1211.

[28] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, P. A. M. K. Kanda, R. Nitrini, and R. Anghinah, "Does EEG Montage Influence Alzheimer's Disease Electroclinic Diagnosis?" *International Journal of Alzheimer's Disease*, vol. 2011, pp. 1–6, 2011.

[29] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, P. A. Kanda, R. Anghinah, and R. Nitrini, "Improving Alzheimer's Disease Diagnosis with Machine Learning Techniques," *Clinical EEG and Neuroscience*, vol. 42, no. 3, pp. 160–165, Jul. 2011.

[30] R. Cassani, T. H. Falk, F. J. Fraga, P. A. M. Kanda, and R. Anghinah, "The effects of automated artifact removal algorithms on electroencephalography-based Alzheimer's disease diagnosis," *Frontiers in Aging Neuroscience*, vol. 6, p. 55, 2014.

[31] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *Journal of Neuroscience Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006.

[32] M. T. Akhtar, W. Mitsuhashi, and C. J. James, "Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data," *Signal processing*, vol. 92, no. 2, pp. 401–416, 2012.

[33] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.

[34] L. Atlas and S. A. Shamma, "Joint acoustic and modulation frequency," *EURASIP Journal on Advances in Signal Processing*, vol. 2003, no. 7, p. 310290, 2003.

[35] R. Cassani and T. H. Falk, "Spectrotemporal Modeling of Biomedical Signals: Theoretical Foundation and Applications," in *Reference Module in Biomedical Sciences*. Elsevier, 2018.

[36] D. P. Tobon V., T. H. Falk, and M. Maier, "MS-QI: A Modulation Spectrum-Based ECG Quality Index for Telehealth Applications," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1613–1622, Aug. 2016.

[37] C. S. Herrmann, M. Grigutsch, and N. A. Busch, "EEG oscillations and wavelet analysis," *Event-related potentials: A methods handbook*, p. 229, 2005.

[38] V. Bewick, L. Cheek, and J. Ball, "Statistics review 13: Receiver operating characteristic curves," *Critical care*, vol. 8, no. 6, p. 5, 2004.

[39] R. Anghinah, P. A. M. Kanda *et al.*, "Alzheimer's disease qEEG: Spectral analysis versus coherence. which is the best measurement?" *Arquivos de neuro-psiquiatria*, vol. 69, no. 6, pp. 871–874, 2011.

[40] F. Pedregosa, G. Varoquaux *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

[41] R. D. Pascual-Marqui, "Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: Exact, zero error localization," *arXiv preprint arXiv:0710.3341*, 2007.

[42] J. Mazziotta, A. Toga *et al.*, "A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM)," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 356, no. 1412, pp. 1293–1322, Aug. 2001.

[43] M. Brett, I. S. Johnsrude, and A. M. Owen, "The problem of functional localization in the human brain," *Nature Reviews Neuroscience*, vol. 3, no. 3, pp. 243–249, Mar. 2002.

[44] J. N. Ianof, F. J. Fraga *et al.*, "Comparative analysis of the electroencephalogram in patients with Alzheimer's disease, diffuse axonal injury patients and healthy controls using LORETA analysis," *Dementia & Neuropsychologia*, vol. 11, no. 2, pp. 176–185, Jun. 2017.

[45] A. Majkowski, Ł. Oskwarek, M. Kołodziej, and R. J. Rak, "An attempt to localize brain electrical activity sources using EEG with limited number of electrodes," *Biocybernetics and Biomedical Engineering*, vol. 36, no. 4, pp. 686–696, 2016.

[46] R. D. Pascual-Marqui, "Review of Methods for Solving the EEG Inverse Problem," vol. 1, no. 1, p. 13, 1999.

[47] J.-S. Kim, S.-H. Lee *et al.*, "Clinical Implications of Quantitative Electroencephalography and Current Source Density in Patients with Alzheimer's Disease," *Brain Topography*, vol. 25, no. 4, pp. 461–474, Oct. 2012.

[48] R. Lizio, C. Del Percio *et al.*, "Neurophysiological Assessment of Alzheimer's Disease Individuals by a Single Electroencephalographic Marker," *Journal of Alzheimer's Disease*, vol. 49, no. 1, pp. 159–177, Sep. 2015.

[49] F. Vecchio, F. Miraglia *et al.*, "Human brain networks in cognitive decline: A graph theoretical analysis of cortical connectivity from EEG data," *Journal of Alzheimer's Disease*, vol. 41, no. 1, pp. 113–127, 2014.

[50] L. R. Gianotti, G. Künig *et al.*, "Correlation between disease severity and brain electric LORETA tomography in Alzheimer's disease," *Clinical Neurophysiology*, vol. 118, no. 1, pp. 186–196, Jan. 2007.

[51] C. Babiloni, A. I. Triggiani *et al.*, "Classification of Single Normal and Alzheimer's Disease Individuals from Cortical Sources of Resting State EEG Rhythms," *Frontiers in Neuroscience*, vol. 10, Feb. 2016.

[52] C. Babiloni, F. Carducci *et al.*, "Resting state cortical electroencephalographic rhythms are related to gray matter volume in subjects with mild cognitive impairment and Alzheimer's disease," *Human Brain Mapping*, vol. 34, no. 6, pp. 1427–1446, Jun. 2013.

[53] C. Babiloni, G. Frisoni *et al.*, "Hippocampal volume and cortical sources of EEG alpha rhythms in mild cognitive impairment and Alzheimer disease," *NeuroImage*, vol. 44, no. 1, pp. 123–135, Jan. 2009.

[54] S. Rathore, M. Habes, M. A. Iftikhar, A. Shacklett, and C. Davatzikos, "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages," *NeuroImage*, vol. 155, pp. 530–548, Jul. 2017.

[55] B. C. Dickerson, A. Bakkour *et al.*, "The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and is Detectable in Asymptomatic Amyloid-Positive Individuals," *Cerebral Cortex*, vol. 19, no. 3, pp. 497–510, Mar. 2009.

[56] R. Bajo, F. Maestú *et al.*, "Functional connectivity in mild cognitive impairment during a memory task: implications for the disconnection hypothesis," *Journal of Alzheimer's Disease*, vol. 22, no. 1, pp. 183–193, 2010.

[57] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nature Reviews Neurology*, vol. 6, no. 2, pp. 67–77, Feb. 2010.

[58] C. Babiloni, G. Frisoni *et al.*, "Frontal white matter volume and delta EEG sources negatively correlate in awake subjects with mild cognitive impairment and Alzheimer's disease," *Clinical Neurophysiology*, vol. 117, no. 5, pp. 1113–1129, May 2006.

[59] D. Moretti, D. Paternicò, G. Binetti, O. Zanetti, and G. Frisoni, "EEG markers are associated to gray matter changes in thalamus and basal ganglia in subjects with mild cognitive impairment," *NeuroImage*, vol. 60, no. 1, pp. 489–496, Mar. 2012.

[60] E. S. Sharp and M. Gatz, "Relationship Between Education and Dementia: An Updated Systematic Review," *Alzheimer Disease & Associated Disorders*, vol. 25, no. 4, pp. 289–304, 2011.

[61] P. Satz, "Brain Reserve Capacity on Symptom Onset After Brain Injury: A Formulation and Review of Evidence for Threshold Theory," *Neuropsychology*, vol. 7, no. 3, pp. 273–295, 1993.

[62] R. S. Wilson, L. Yu, M. Lamar, J. A. Schneider, P. A. Boyle, and D. A. Bennett, "Education and cognitive reserve in old age," *Neurology*, vol. 92, no. 10, pp. e1041–e1050, 2019. [Online]. Available: https://n.neurology.org/content/92/10/e1041