Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 100 (2016) 865 – 871

Procedia Computer Science

Conference on ENTERprise Information Systems / International Conference on Project MANagement / Conference on Health and Social Care Information Systems and Technologies, CENTERIS / ProjMAN / HCist 2016, October 5-7, 2016

Alzheimer's Early Prediction with Electroencephalogram

Pedro Miguel Rodrigues[a,*], João Paulo Teixeira[b], Carolina Garrett[c], Dílio Alves[c] and Diamantino Freitas[a]

[a]*Faculty of Engineering, University of Porto, Porto, Portugal*
[b]*Politechinic Institute of Bragança, Bragança, Portugal*
[c]*Hospital de São João, Porto, Portugal*

Abstract

Alzheimer's disease (AD) is currently an incurable illness that causes dementia and patient's condition is progressively worse and it represents one of the greatest public health challenges worldwide. The main objective of this work was to develop a classification methodology for EEG signals to improve discrimination amongst patients at varying stages of the illness, Mild Cognitive Impairment (MCI) patients and non-patients either in order to obtain a more reliable methodology to identify AD in early stages. For this purpose, a surrogate decision tree classifier was used with 2 different ways of cross-validation (leave-one-out-cross-validation and 10-fold-cross validation). The EEG studied features were the values of maxima (NMax) and minima (NMin), the zero-crossing (Zcr) rate, the mean derivative value at a point (Mdif), the Relative Power (RP) in each of the conventional bands and finally the spectral ratios (r). The best classification was obtained with vectors of 10 features as classifier entries in a leave-one-out-cross validation process, reaching 0.934 AUC, a sensitivity of 86.19%, a specificity of 99.35%, an accuracy of 94.88%, with a low out-of-sample classification error of 6.98% which indicates that the classifier generalizes fairly well.

Peer-review under responsibility of the organizing committee of CENTERIS 2016

Keywords: Alzheimer's diasease; Early stages; Electroencephalogram signal; surrogate decision tree classifier; features;

* Corresponding author. Tel.: Tel.: +351 96 891 67 85; fax: +351 22 508 14 40.
E-mail address: pedro.luis.rodrigues@fe.up.pt

Peer-review under responsibility of the organizing committee of CENTERIS 2016
doi:10.1016/j.procs.2016.09.236

1. Introduction

Alzheimer's disease (AD) is a progressive and irreversible neurodegenerative brain disorder and is currently the most common cause of dementia in the elderly[1].

This disease is an acquired disorder of cognitive and behavioral impairment that notoriously interferes with social or occupational functioning. The cause of AD is not yet known at this time[2]. The progression of the disease can be classified in four different stages. The first stage or the pré-dementia stage is known as Mild Cognitive Impairment (MCI). The MCI is known as a transitional stage between natural aging and AD. It corresponds to a variety of symptoms: difficulty in remembering recent events, subtle changes in behavior, discrete loss of autonomy in daily life activities, disorientation in time and space and personality changes. Even though the MCI confers an increased risk of AD developing, this state can lead to several possible outcomes, including the improvement to a normal cognitive state. The Mild and Moderate AD stages are characterized by increasing cognitive deficits and complete dependence on caregivers. In the Advanced stage monitoring becomes constant and strictly necessary because patients are unable to perform any tasks and so assistance is inevitable[3].

The early diagnostic accuracy is low and there is not a biomarker able to detect AD without invasive tests. An autopsy or brain biopsy is the only way to make a definitive diagnosis. Effective diagnosis can provide opportunities for AD patients to get involved in clinical trials and to get the best treatment[4].

In the last years, many research groups have started investigating the potential of electroencephalograms (EEGs) for AD diagnosis. For several decades the EEG has been used as a diagnostic tool for dementia as it is a very useful tool in the study of brain disorders being a non-invasive technique that records the electromagnetic fields produced by the brain activity[4].

The EEG is a noninvasive technique that records the electromagnetic fields produced by brain activity with good temporal resolution. EEG is typically divided in different conventional frequency bands, such as delta (δ, 1-4 Hz), theta (θ, 4-8 Hz), alpha (α, 8-13Hz), beta (β, 13-30 Hz), and gamma (γ, 30-40 Hz). AD seems to affect the signal power in those different bands. The major effect is known as the EEG "slowing", that means a power increase in low frequency bands such as delta and theta, and a power decrease in higher frequency bands such as alpha and beta[4].

In this study the authors intend to develop a classification methodology that allows to distinguish AD patients in Mild/Moderate and advanced stages, MCI patients and Control Subjects in order to get as far as possible a more reliable AD diagnosis in early stages. For that, a surrogate decision tree classier with leave-one-out-cross-validation and with 10-fold-cross-validation will be used. This paper is organised as follows: section Introduction – this section focuses on the State of Art and at the study main objective; section Materials – this section presents the study participants and the EEG recording; Proposed Methodology - in this section the methodologies employed are described, finally, the section Discussion and Conclusion show the discussion of the results achieved and concludes the paper emphasizing its main contributions.

2. Materials

A set of 37 subjects participated in this study (11 as Controls, 8 with MCI, 10 with AD in Mild/Moderate stages and 8 in Advanced stage). EEGs were recorded from the 19 scalp loci of the international 10-20 configuration using a digital electroencephalograph in "Hospital de São João - Porto", Portugal. The sampling frequency was 256 Hz and all recordings were digitally filtered with a band-pass filter with cut-off frequencies at 1-40Hz. In the next table more information, such as the Age average of the study participants in each group and the average Mini Mental State Examination (MMSE) index achieved in each group, can be found. All EEG were organized in segments of 5s (1280 epochs).

Table 1. *The Dataset*

Dataset				
	Control Subjects	**MCI Patients**	**AD Mild/Moderate stages**	**AD Advanced stage**
#	11	8	10	8
Age average	74	80	79	79
MMSE average	28.68	26.29	18.89	11.50

3. Proposed Methodology

The proposed methodology is composed by five different steps. The first three steps are related to the EEG signal processing and features extraction and the last two about statistical analysis and classification. The methodology steps organization is illustrated in Fig. 1.

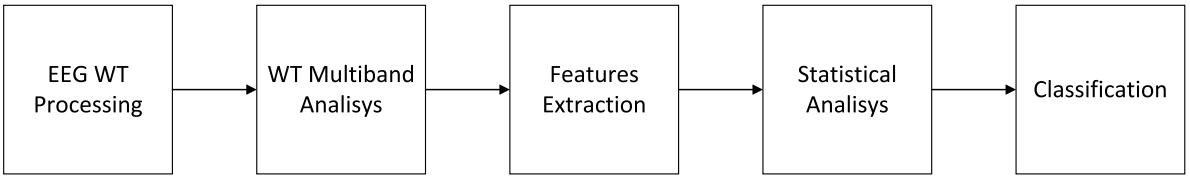

Fig. 1. Methodology description

3.1. EEG Wavelet Transform Processing

The Wavelet Transform (*WT*) was chosen to be the main technique to describe the EEG signals. This signal processing tool provides a good time resolution for the high frequencies and good frequency resolution for low frequencies, which is useful for processing EEG signals because the target frequencies bands were localized in low frequencies.

The Discrete Wavelet Transform (*DWT*) was used in the implementation. Its multiresolution analysis characteristics consist in obtaining increasingly finer signal resolution versions by successive filtering. The *DWT* uses a two functions set: a scale function ($\varphi[n]$) and a Wavelet function ($\psi[n]$)[5], as follows:

$$\varphi[n] = \sum_k h[k] \cdot \phi[2 \cdot n - k] \tag{1}$$

$$\psi[n] = \sum_k g[k] \cdot \phi[2 \cdot n - k] \tag{2}$$

where k is the discrete translation parameter and $h[k]$ and $g[k]$ are, respectively, the impulse responses of the low pass and high pass filters used in the *WT* analysis. The signal decomposition in different frequency bands is achieved by successive low-pass filters and high-pass filters in the time domain, followed by subsampling by a factor of 2 until the maximum level of decomposition ($log_2(N)$) is reached[5] as is illustrated in the following equations 3 and 4:

$$DWT_A^{\varphi}[j,k] = \sum_{n=0}^{\frac{N}{2^{(j-1)}}} DWT_A^{\varphi}[j-1,n] \cdot h[2 \cdot k - n] \tag{3}$$

$$DWT_D^{\psi}[j,k] = \sum_{n=0}^{\frac{N}{2^{(j-1)}}} DWT_A^{\psi}[j-1,n] \cdot g[2 \cdot k - n] \tag{4}$$

where, k=0, ..., N/(2^j); j=0, ..., $log_2(N)$, A represents the *DWT* approximated coefficients, D the *DWT* detail coefficients and N the signal length[4]. The Power Spectral Density (*PSD*) functions showed the variations strength of Power as a function of frequency. The Wavelet Packet Transform (*WPT*) spectrum was obtained by extracting the *WPT* coefficients corresponding to the terminal nodes. The *WPT* function was defined as:

$$X_{\varphi,\psi}[j,m,k] = \sum_{n=0}^{\frac{N}{2^j-1}} WP^{\varphi,\psi}[j,m,n] \cdot w_{j,m,k}[n],$$

$$w_{j,m,k}[n] = 2^{-\frac{j}{2}} \cdot w_m\left[2^{-j} \cdot n - k\right],$$

$$j = 1,\ldots,J;\ m = 0,1,\ldots,2^j - 1;\ k = 1,\ldots,2^{J-j} \quad (5)$$

where $J=log_2(n)$, j represented the level of *WPT*. To estimate the *PSD* of *the WT*, it was used the *WPT* periodogram based on the observations of length $N=2^j-1$.

$$S_{x,\varphi,\psi}[m,k] = \left[\sum_{j=1}^{J} X_{\varphi,\psi}[m,k]\right]^2;\ m = 0,1,\ldots,2^j - 1;\ k = 1,\ldots,2^{J-j}; \quad (6)$$

To obtain the frequency spacing of the *PSD* of *the WT* of each level of decomposition it was used the range,

$$\left[\frac{F_s}{2^{j+1}}, \frac{F_s}{2^{j+1-1}}\right],\quad j = 1,\ldots,log_2(N), \quad (7)$$

Summarizing, the absolute value of the coefficients was taken and the Wavelet coefficients were ordered by frequency. The *PSD* was normalized (PSD_n) as below[6,7]:

$$PSD_n[m,k] = \frac{S_{x,\varphi,\psi}[n]}{\sum_{n=k_1}^{k_2} S_{x,\varphi,\psi}[n]},\quad m = 0,\ldots,2^j - 1 \quad (8)$$

3.2. WT Multiband Analysis

In order to extract some information from the EEG signals of AD patients, in Mild/Moderate and Advanced stages, MCI patients and Control Subjects, each signal resulting from the overall average process was decomposed to the level 5 by the Wavelet Biorthogonal (Bior) 3.5 *DWT*. Additionally the EEG signals were processed by Bior 3.5 *PSD* of *WT* technique to the level 8 of decomposition[6-9]. The wavelet Bior3.5 was used to process the EEG signals because it had more discriminant probabilities evaluated in previous works and it proved to be a good choose in EEG signal processing[6-9]. These decomposition levels for both techniques provided the correct level of EEG signal decomposition to reach the conventional frequency bands of EEG (δ, θ, α, β and γ bands). In the case of the *PSD* of WT technique, the level 8 of decomposition allowed to obtain a *PSD* which ranges between 1Hz and 128Hz. Thus containing the frequency range of conventional EEG bands (1-40Hz) and in this way, the *PSD* of WT can be restricted to this range of frequencies. In the case of DWT, the level 5 of decomposition gives the detail coefficients D2 that corresponding to the Gamma band, D3 to the beta band, D4 to the alpha band, D5 to the theta band and the approximated coefficients A5 to the delta band.

Afterwards several features were extracted in the relevant frequency bands (δ, θ, α, β and γ) of the EEG signal. These features are intended to characterize the waveform in terms of its frequency content and of greater or smaller variability as is required to observe the "*shift-to-the-left*" phenomenon. Selected features were the values of maxima (***NMax***) and minima (***NMin***), the zero-crossing (***Zcr***) rate, the mean derivative value at a point (***Mdif***), the Relative Power (***RP***) and the spectrum energy deceleration represented by the spectral ratios (***r***), as indicated from the previous studies[8,9]:

- ***NMax*** and ***NMin***: Accounting and calculation of all signal maxima and minima, respectively, by the variation of the signal waveform derivative[8,9];
- ***Zcr*****:** The zero-crossing rate was calculated as[8,9]:

$$Zcr_i = \frac{1}{N-1}\sum_{n=1}^{N-1} \mathrm{I}\left\{s_i(n)\,s_i(n) - 1 < 0\right\}, \quad (9)$$

where $i = \{\delta, \theta, \alpha, \beta, \gamma\}$, s is a *DWT* coefficient signal of length N and I the indicator function. If the I value, in a certain position, is equal to 1 a signal zero crossing is found.

- ***Mdif*:** The mean derivative value at a point was obtained by the equation 5[8,9],

$$Mdif_i = \frac{1}{N-6} \cdot \sum_{n=1}^{N-6} \left(s_i(n) - \sum_{b=n}^{n+5} s_i(b) \right), \quad (10)$$

where $i = \{\delta, \theta, \alpha, \beta, \gamma\}$.

- ***RP*:** It was calculated as the sum of the *PSD* components in the conventional frequency bands: delta (1-4 Hz, δ), theta (4-8 Hz, θ), alpha (8-13Hz, α), beta1 (13-19Hz, $\beta 1$), beta2 (19-30 Hz, $\beta 2$) and gamma (30-40 Hz, γ)[6-8];

- ***r***: Four spectral ratios were used to describe the deceleration of the EEG spectrum of AD patients. r_1 is a spectral ratio that enabled us to detect changes when a slight slowdown in the EEG spectrum appears. The r_2 index summarizes the EEG globally slowing down. r_3 computes the relationship between high frequency bands ($\beta 1$ and $\beta 2$) and the lowest frequency band (δ). r_4 is similar to r_3, but with a narrower high frequency band (only $\beta 2$)[6-9].

$$r_1 = \frac{RP(\alpha)}{RP(\theta)}; \quad (11)$$

$$r_2 = \frac{RP(\alpha) + RP(\beta 1) + RP(\beta 2) + RP(\gamma)}{RP(\delta) + RP(\theta)}; \quad (12)$$

$$r_3 = \frac{RP(\beta 1) + RP(\beta 2)}{RP(\delta)}; \quad (13)$$

$$r_4 = \frac{RP(\beta 2)}{RP(\delta)}; \quad (14)$$

3.3. Statistical Analysis

The normality of the extracted parameters was assessed with the Kolmogorov-Smirnov test and the homoscedasticity with the Levene's test. Data distributions did not meet the hypotheses of parametric tests. Thus, differences between groups were analyzed using the Mann-Whitney U test (p<0.05). In Table 2 the best p-values achieved for each feature are represented in bold.

Table 2. Differences between extracted features distributions of study groups

Features	*p-value*	Features	*p-value*	Features	*p-value*	Features	*p-value*	Features	*p-value*	Features	*p-value*
$NMax_\delta$	0,1404	$NMin_\delta$	0,1411	Zcr_δ	**0,0049**	$Mdif_\delta$	**0,0310**	E_δ	**0,0075**	r_1	**0,0025**
$NMax_\theta$	**0,0023**	$NMin_\theta$	**0,0023**	Zcr_θ	**0,0049**	$Mdif_\theta$	**0,0379**	E_θ	**0,0050**	r_2	**0,0011**
$NMax_\alpha$	**0,0224**	$NMin_\alpha$	**0,0224**	Zcr_α	**0,0049**	$Mdif_\alpha$	**0,0408**	E_α	**0,0049**	r_3	**0,0147**
$NMax_\beta$	0,9012	$NMin_\beta$	0,8422	Zcr_β	**0,0049**	$Mdif_\beta$	**0,0163**	E_β	**0,0463**		
$NMax_\gamma$	0,9012	$NMin_\gamma$	0,8422	Zcr_γ	**0,0049**	$Mdif_\gamma$	**0,0163**	E_γ	0,6976		

Note: In bold p<0.05

3.4 Classification

Concerning the EEG processing at this step the objective was targeted with a *leave-one-out-cross-validation* (LCV) and *10- fold-cross-validation* (10FCV) classification of EEG study signals. For getting more data to generalize the classification an average over the channel was performed for each one of the study participants and at the end 19

different signals were achieved from each one. From the new signals, the previously introduced features at point 3.2 were extracted to serve as an individual classifier vector entry. The features that did not provide a $p<0.05$ in previous section were not used at this step. In each one of different kind of cross validations, for the classifier that used features extracted from a study participant channel signal, the others 18 vectors of features extracted from the other channels of the same participant were not used in the classification, in order to avoid overfitting. A genetic algorithm was used to get the best features combinations between 5 to 12 features that served as vectors entries for the classifier. The chosen classifier was surrogate decision trees and it was evaluated by using Receiver Operating Characteristic (*ROC*) curves. The classification results can be seen in next Table 3. In bold, in Table 3, the best classifications achieved are shown, and the one that produced the lowest out-of-sample classification is highlighted, this value (6,98%) indicating that the classifier generalizes fairly well.

Table 3. *Classification values for both methods of classification*

Classification Type	# of features	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)	The out-of-sample classification error (%)
LCV	5	0.9168	83.97%	97.85%	93.17%	10.80%
10FCV		0.9316	86.96%	98.10%	94.45%	10.19%
LCV	6	0.9421	88.79%	97.71%	94.88%	11.28%
10FCV		0.9424	88.74%	97.51%	94.74%	9.82%
LCV	7	0.9228	86.34%	97.27%	93.74%	10.80%
10FCV		0.9392	88.99%	98.53%	95.45%	9.32%
LCV	8	0.9357	88.79%	97.71%	94.88%	11.78%
10FCV		0.9398	88.89%	98.12%	95.16%	10.81%
LCV	9	0.9356	86.13%	99.14%	94.74%	8.64%
10FCV		0.9272	87.95%	97.49%	94.45%	10.61%
LCV	**10**	**0.9339**	**86.19%**	**99.35%**	**94.88%**	**6.98%**
10FCV		0.9171	84.32%	97.86%	93.31%	10.60%
LCV	11	0.9336	85.83%	99.35%	94.74%	9.14%
10FCV		0.9144	83.47%	98.48%	93.31%	11.30%
LCV	12	0.9072	81.93%	98.90%	92.89%	10.50%
10FCV		0.9135	86.43%	96.27%	93.17%	12.61%

4. Discussion and Conclusion

Many studies have been done about AD but what causes this disease is not fully understood yet. If AD is detected in early stages appropriate treatments may delay its effects and can help preserve daily functioning at least for a longer time once the disease progression will continue because there is no known cure for AD. But if correct diagnosis is achieved only in an advanced stage of AD the situation is critical and almost nothing can be done to help the patient, that is why the authors feel more and more motivated and encouraged to develop their work in order to give a useful contribution for AD early diagnosis. Moreover, the present study is based on EEG because the authors believe that it has several good characteristics and low-cost and therefore it appears to be a good objective to increase its value as a diagnosis tool. In short, the authors of this study processed the global average of each EEG signal of AD patients and control subjects with the *WT* to get each reconstructed conventional band, from that signal, respectively. After that, several features had been extracted. Significant differences were found between the features distributions, respectively, of the study groups. The best features were $NMax_{\theta}$, $NMin_{\theta}$, all *Zr* in all conventional bands, E_{δ}, E_{θ}, E_{α} and the index r_1 *and* r_2, but other features were found that also provided $p<0.05$. This step was done to optimize and find good features that, when combined, may increase the accuracy of the classifier. After this an average over the channel was performed for each one of the study participants and previously selected features with a $p<0.05$ were extracted. Afterwards a surrogate decision trees classifier with *LCV* and with *10FCV* was built and the best classification achieved was 0.934 of AUC, 86.19% of sensitivity, 99.35% of specificity and 94.88% of accuracy, with a low out-of-sample classification error of 6.98% which indicates that the classifier generalizes fairly well. Several studies tried to discriminate between controls and AD, AD and MCI and MCI and controls but only few tried to discriminate between AD, MCI and controls subjects. Comparatively with those few studies where the better accuracy

achieved was 77%[10], in this study the classifier provided a substantially higher accuracy of 94.88%. In spite of that the results should be extended on a larger population to ensure generalization.

With this study it was revealed that a surrogate decision tree classifier can be a good tool for EEG signal processing in order to help in AD early identification.

References

1. Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. Alzheimer's disease. The Lancet 2011; **377**:1019-1031.
2. Jonker C, Launer L, Hooijer C, Lindeboom J. Memory complaints and memory impairment in older individuals. J Am Geriatr Soc 1996; **44**:44-49.
3. Nygard L. Instrumental activities of daily living: a stepping-stone towards Alzheimer's disease diagnosis in subjects with mild cognitive. Acta Neurol Scand. 2003, **179**: 42-48.
4. Jeong J. Eeg dynamics in patients with alzheimer's disease. Clin. Neurophysiol 2004; **115**:1490-1505.
5. Rioul O, Vetterli, M. Wavelets and signal processing. IEEE Signal Processing Magazine 1992; **8**:14-38.
6. Rodrigues P.M., Teixeira J.P. Alzheimer's Disease Recognition with Artificial Neural Networks, In Information Systems and Technologies for Enhancing Health and Social Care, by R. Martinho, R. Rijo, M. Cruz-Cunha and J. Varajão (Eds.), IGI Global 2013,102-118, 2013.
7. Rodrigues P, Teixeira J. Artificial neural networks in the discrimination of alzheimer's disease. Communications in Computer and Information Science 2011; **221**:272-281.
8. Rodrigues, P. M., Freitas, D. R. and Teixeira, J.P. Detection of Alzheimer's Disease Electroencephalogram Temporal Events. International Journal of Reliable and Quality E-Healthcare (IJRQEH) 2013, **2**:44-61.
9. Rodrigues, P. M., Freitas, D. R. and Teixeira, J.P. Alzheimer's electroencephalogram event scalp localization", IEEE 9th International Workshop on Multidimensional (nD) Systems (nDS); Vila Real, 2015, 1-4.
10. Akrofi, K., Pal, R., Baker, R.C., Nutter, B. S. and Schiffer, R. W. Classification of Alzheimer's disease and mild cognitive impairment by pattern recognition of EEG power and coherence, *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, Dallas, 2010, 606-609.