TOWARDS AN EEG-BASED BIOMARKER FOR ALZHEIMER'S DISEASE: IMPROVING AMPLITUDE MODULATION ANALYSIS FEATURES

Francisco J. Fraga[1,2], Tiago H. Falk[1], Lucas R. Trambaiolli[2], Eliezyer F. Oliveira[2], Walter H. L. Pinaya[2], Paulo A. M. Kanda[3], Renato Anghinah[3]

[1]INRS-EMT, University of Quebec, Montréal, Quebec, Canada
[2]Universidade Federal do ABC (UFABC), CECS, Santo André, Brazil
[3]Universidade de São Paulo (USP), School of Medicine, CEREDIC, São Paulo, Brazil

ABSTRACT

In this paper, an EEG-based biomarker for automated Alzheimer's disease (AD) diagnosis is described, based on extending a recently-proposed "percentage modulation energy" (PME) metric. More specifically, to improve the signal-to-noise ratio of the EEG signal, PME features were averaged over different durations prior to classification. Additionally, two variants of the PME features were developed: the "percentage raw energy" (PRE) and the "percentage envelope energy" (PEE). Experimental results on a dataset of 88 participants (35 controls, 31 with mild-AD and 22 with moderate AD) show that over 98% accuracy can be achieved with a support vector classifier when discriminating between healthy and mild AD patients, thus significantly outperforming the original PME biomarker. Moreover, the proposed system can achieve over 94% accuracy when discriminating between mild and moderate AD, thus opening doors for very early diagnosis.

Index Terms — EEG-based biomarker, Alzheimer's disease diagnosis, amplitude modulation analysis

1. INTRODUCTION

After cardiovascular disease and cancer, Alzheimer's disease (AD) is now the third most expensive disease and the sixth leading cause of death in the United States [1,2]. Nevertheless, there is not yet a precise biomarker to define AD and definitive diagnosis can only be established with a hystopathological analysis of the brain [3]. Hence, AD diagnosis is done based upon clinical history, laboratory tests, neuroimaging and neuropsychological evaluations, with accuracies ranging from 85-93% [4]. However, these clinical assessments are nonspecific and costly, and require experienced clinicians and lengthy sessions [5].

As a consequence, there is an urgent need for an accurate, universal, specific and cost-effective biomarker to diagnose AD and to follow disease progression and therapy response. Early diagnosis is crucial in order to initiate treatment that can retard disease progression. Over the last decade, there has been an effort to develop computer-based tools to assist physicians in making more precise and earlier diagnostics. Towards this goal, quantitative electro-encephalogram (qEEG) has emerged as a promising tool [6].

Previous findings have suggested that i) EEG spectral power is reduced with AD in the alpha (8-12 Hz) and beta (12-30 Hz) frequency bands, and increased in the delta (0.1-4 Hz) and theta (4-8 Hz) bands [7], ii) spectral coherence is decreased between the two hemispheres in the alpha and beta frequency bands [8], and iii) EEG pattern complexity [9–11] is reduced. More recently, a new promising biomarker was developed, which was termed "percentage modulation energy" (PME) [12]. The marker characterizes the amplitude modulation rate-of-change of resting-awake EEG signals and showed significant differences in the amplitude modulation of theta and beta frequency bands between AD patients and controls.

To test the effectiveness of the PME marker (henceforth referred to as a "feature") in discriminating AD from healthy controls, a support vector machine classifier was developed and shown to achieve over 90% accuracy, sensitivity, and specificity [13]. While these results were promising on their own, they still remained below the best rates achieved by experienced clinicians [4]. In an effort to improve recognition performance, this paper proposes two updates to the PME-based diagnostic system. First, in order to improve the signal-to-noise ratio (SNR) of the EEG signal, we propose to average PME features over several epochs prior to classification. This is akin to the averaging done for calculation of EEG evoked potentials [14]. Previously, PME features were computed on a per-epoch basis [13], i.e., without averaging. Second, we propose two additional features termed "percentage raw energy" (PRE) and "percentage envelope energy" (PEE) to characterize overall per-band amplitude modulation rate-of-change.

On a dataset of 88 participants (35 controls, 31 with mild-AD and 22 with moderate AD), the newly proposed features are shown to outperform the original PME when discriminating AD from controls, and mild from moderate AD,thus opening doors for very early diagnosis.

[1] The first author performed the work while a Visiting Scholar at INRS and thanks to Brazilian agency FAPESP for funding.

2. MATERIALS AND METHODS

2.1. Subjects

In this study, AD patients and control subjects 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 Clínicas in São Paulo, Brazil. Diagnosis was made by experienced neurologists based on the Brazilian version of the Mini-Mental State Examination (MMSE)[15] and the Clinical Dementia Rating (CDR) scale.

Resting-awake multi-channel EEG recordings were obtained from 88 participants, separated into three education-matched groups. The first group was composed by 35 normal subjects (NS), 19 females and 16 males (mean age 66.9 years, 8.2 sd). Inclusion criteria for cognitively normal cohort were CDR score = 0, MMSE score $\geq$ 25 (mean 28.0, 2.2 sd) and no indication of functional cognitive decline prior to enrollment based on an interview with the subject. The second group (AD1) comprised 31 mild-AD patients (according to NINCDS-ADRDA [16] and DSM-IV-TR [17] criteria), 19 females and 12 males (mean age 75.2 years, 5.6 sd). Other inclusion criteria for AD1 group were $0.5 \leq$ CDR ≤ 1 and MMSE ≤ 24 (mean 19.5.0, 3.6 sd). The third group (AD2) included 22 AD patients with moderate AD (DSM-IV-TR), 15 females and 7 males (mean age 73.8 years, 10.2 sd). Inclusion criteria for AD2 group were CDR score = 2 and MMSE score $\leq$ 20 (mean 14.2.0, 3.7 sd).

For inclusion in both AD cohorts (AD1 and AD2) an additional criterion was the presence of functional and cognitive decline over the previous 12 months based on detailed interview with a knowledgeable informant. Patients from both AD groups were also screened for diabetes mellitus, kidney disease, thyroid disease, alcoholism, liver disease, lung disease or vitamin B12 deficiency, as these can causes cognitive decline. Ethics approval was obtained from the institutes and participants consented to the study.

2.2. Data Acquisition and pre-processing

EEGs were recorded with 12 bits resolution and sampling rate of 200 Hz using the Braintech 3.0 instrumentation (EMSA Equipamentos Médicos Inc., Brazil). Impedance was maintained below 10 kΩ, and the electrodes (referential montage) were placed according to the International 10–20 System. Biauricular referential electrodes were attached as recommended by the Brazilian Society of Clinical Neurophysiology and the American EEG Society.

Low-pass filtering was accomplished using an infinite impulse response elliptic filter with a zero at 60 Hz, to ensure elimination of any power grid interference. From the referential montage we derived a virtual interhemispheric bipolar montage, as there is evidence of an interhemispheric disconnection in AD [18,19,12]. The so-called "bipolar signal" was obtained by simply subtracting the two bi-auricular referenced signals involved [20]. For this work, the bipolar signals were Fp1-Fp2, F3–F4, F7–F8, C3–C4, T3–T4, P3–P4, T5–T6, and O1–O2. During examination, probands were awake and relaxed, with closed eyes. Two skilled neurophysiologists manually removed EEG artifacts (e.g., blinking, muscle movements) from the recordings. Subsequently, from each EEG signal, 28 epochs of eight seconds were selected by visual inspection.

2.3. EEG amplitude modulation analysis

As mentioned previously, an innovative EEG amplitude modulation feature termed "percentage modulation energy" (PME) was proposed for automated AD diagnosis. While a complete description of the feature is beyond the scope of this paper, we provide a brief overview for the sake of completeness. The interested reader is referred to [12,13] for complete details.

In order to compute the so-called PME features, first elliptic bandpass filters are used to decompose the full-band EEG signal into the five well-known subbands: delta, theta, alpha, beta, and gamma [21]. The temporal amplitude envelope $e_i(n)$ of each of the five subband EEG signals $s_i(n)$ is then computed by means of a Hilbert transform. With the aim of directly quantifying the rate-of-change of the subband temporal envelopes and to uncover possible cross-frequency interactions, temporal envelopes are further decomposed into four modulation frequency bands, henceforth referred to as m-delta (0.1–4 Hz), m-theta (4–8 Hz), m-alpha (8–12 Hz), and m-beta (12–30 Hz) using 2nd order bandpass filters. The m-gamma modulation band is not used, motivated by previous findings [13]. By taking a short-term discrete Fourier transform of the doubly-decomposed (framed) signals, the per-frame amplitude modulation energy $E_{i,j}(k)$ is obtained, where 'i' indexes the filters of the first decomposition, 'j' the filters of the second "modulation domain" decomposition for frame 'k'. As such, $\bar{E}_{i,j}$ denotes the average modulation energy over all frames in the 8s-epoch. The *PME* feature is thus computed by:

$$PME_{i,j} = \frac{\overline{E}_{i,j}}{\sum_{i=1}^{5}\sum_{j=1}^{4}\overline{E}_{i,j}} \times 100\% \qquad (1)$$

for each of the 8 bipolar EEG signals.

2.4. Overall Per-Band Modulation Features

Here, two new "overall" per-band modulation features are investigated, namely the *percentage "raw" energy* (*PRE*) and the *percentage envelope energy* (*PEE*) features, computed for each subband signal, and given by:

$$PRE_i = \frac{\sum_{n=1}^{N} s_i^2(n)}{\sum_{i=1}^{5}\sum_{n=1}^{N} s_i^2(n)} \times 100\% \qquad (2)$$

$$PEE_i = \frac{\sum_{n=1}^{N}\left[e_i(n)-\bar{e}_i\right]^2}{\sum_{i=1}^{5}\sum_{n=1}^{N}\left[e_i(n)-\bar{e}_i\right]^2}\times 100\% \qquad (3)$$

for each of the 8 bipolar EEG signals, where N is the 8s-epoch length (in samples), since for *PRE* and *PEE* computation, framing is not performed. It is important to remark that information contained in *PEE* is not present in *PRE* as DC components $\bar{e}_i$ of temporal envelopes are discarded during the frequency decomposition step. In total, for each participant, 240 (5x4=20 *PME*s + 5 *PRE*s + 5 *PEE*s for 8 bipolar signals) features are extracted for each epoch. Due to Bedrosian's theorem [22], however, not all of these features convey useful information, thus a total of 192 (14 *PME*s + 5 *PRE*s + 5 *PEE*s for 8 bipolar signals) are used.

2.5. SNR improvement during feature extraction

In order to improve the SNR of the extracted features, an averaging procedure is performed, akin to what is done in EEG evoked responses studies. To this effect, we investigate the benefits of averaging over different number of epochs (*Ne* = 0, 5, 10, 15, 20), where *Ne*=0 corresponds to the original PME feature without averaging.

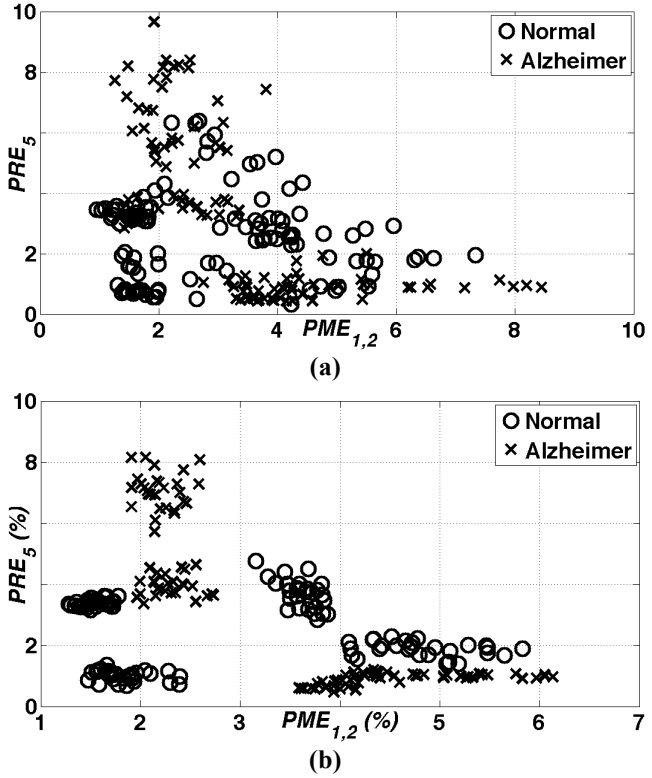

Fig 1. Features from epochs with (a) *Ne = 0* and (b) *Ne* = 5 for the T5-T6 bipolar signal.

Epochs are randomly selected from the pool of 28 available epochs per participant. To better understand the benefits of this SNR improvement stage, the scatter plots in Figures 1 (a) and (b) are used. The plots depict PRE_5 *vs* $PME_{1,2}$ features extracted from the T5-T6 bipolar signal without (*Ne* = 0) and with SNR improvement (*Ne* = 5), respectively. As can be seen, greater feature separability is attained between healthy controls (circles) and AD patients (X) after SNR improvement.

2.6. Feature selection and classifier design

Motivated by previous findings [23,24], we explore the use of support vector machines (SVM) for feature selection and classification. A complete description of these steps is beyond the scope of this paper and the interested reader is referred to [25] and [26] for more details on SVMs. In our experiments, we use only the top-ranked 24 features to remain inline with previous classifiers reported in the literature [24]. When testing the effectiveness of the 'SNR improvement' step, the 24 top-ranked features were selected for each epoch-averaging condition (i.e., *Ne* = 0, 5, 10, 15, and 20). We used the *Weka* polykernel implementation with the following default parameter values: regularization coefficient C = 1 and γ = 0.01. A leave-one-subject-out (LOSO) cross-validation paradigm was performed for classifier design to avoid overfitting and to guarantee the generality of the classifier to unseen data.

3. RESULTS

Using the LOSO paradigm, different classification tests were performed. First, classifier accuracy is reported for three types of two-class discrimination tasks: NS vs AD1, AD1 vs AD2, and NS vs AD; this last task pools AD1 and AD2 into a global AD group. Each of the three two-class discrimination tasks described above are performed for each of the five SNR improvement conditions, i.e., *Ne* = 0, 5, 10, 15, and 20. Table 1 reports the classifier accuracy, sensitivity, and specificity as performance metrics for 'control vs. AD' for different SNR improvement conditions.

Table 1 – Normal vs. Alzheimer classification results (%)

Ne	Accuracy (NxAD)	Specificity (N)	Sensibility (AD)
0	94.32	91.43	96.23
5	96.59	94.29	98.11
10	97.73	100.00	96.23
15	**98.86**	**100.00**	**98.11**
20	96.59	97.14	96.23

Figure 2 shows the classifier accuracy (vertical axis) for five different *Ne* values (horizontal axis) obtained in the three two-class discrimination tasks: NS x AD1 (dotted line), AD1 x AD2 (dashed line) and NS x AD (solid line). It is possible to see in Figure 2 that classification accuracy increased significantly as the number of averaged epochs increases but started decreasing after *Ne* = 15.

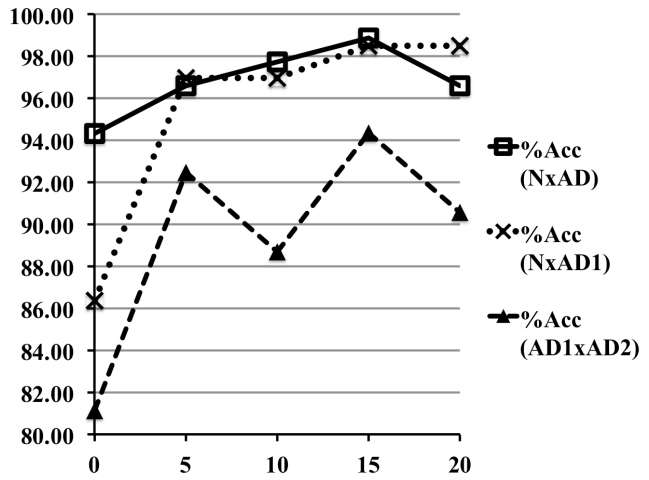

Fig. 2. Classifier accuracy versus number of averaged epochs *Ne* for the three 2-class problems

Both Table 1 and Figure 2 present classifier performance not on a "per-epoch" basis, but on a more medically relevant "per-participant" basis, where a patient is deemed to be correctly classified when the majority of its 28 epochs have been correctly classified.

4. DISCUSSION

4.1. Benefits of SNR improvement

It is interesting to note that SNR improvement was more important for the NS x AD1 and AD1 x AD2 classifications than for the NS x AD case. This is an important advantage, as it allows for the disease to be detected at an early stage (normal vs. mild AD discrimination), as well as for disease progression to be automatically monitored (mild vs. moderate AD discrimination), thus potentially opening doors for in-home telemedicine or tele-rehabilitation tools to be developed. Notwithstanding, the SNR improvement stage also resulted in significant gains relative to the original PME metric for the 'control vs. AD' classification task.

4.2. Relationship to existing biomarkers

In relationship to previous PME-based studies based on a much smaller dataset (32 participants [12,13]), the present experiment has validated the use of modulation energy features for reliable AD diagnosis on a much larger dataset. Moreover, with the proposed updates and innovative features, significant gains have been obtained relative to the original PME feature (see Table 1 and Fig. 2), in terms of accuracy, sensitivity, and specificity. Moreover, the updated biomarkers have also allowed for very early detection and automated disease severity progression assessment.

Relative to other studies, it is difficult to make direct performance comparisons, as different databases have been used, as have different protocols know to affect EEG patterns (e.g., eyes-closed vs eyes-open; rest vs. mental activity, etc.). As such, only indirect broad comparisons with other state-of-the-art techniques can be made. For example, using a parallel nonlinear EEG analysis technique, Buscema et al. [10] obtained 85.98% accuracy in distinguishing mild cognitive impairment and Alzheimer's disease patients (143 probands), a task similar (but not exactly the same) to our mild vs. moderate classification. In the same direction and using almost the same technique, Rossini et al. [11] got 93.46% accuracy to automatically distinguish resting EEG data of 171 normal elderly vs. 115 Mild Cognitive Impairment (MCI) subjects. Again, this is somewhat close to our normal vs. mild discrimination task, as a CDR of 0.5 is one of the conditions for MCI diagnosis [27] and it matches one of the inclusion criteria for our AD1 group (mild). As can be seen, the classification performance results shown herein are somewhat higher performances than these reported in the literature, but with the additional benefit of allowing for accurate 'healthy vs. mild-AD', 'mild-AD vs. moderate-AD', and 'healthy vs. AD' (mild and moderate pooled together) classification.

4.3 Study limitations and ongoing work

The findings reported herein constitute that of a "semi-automated" system, as the artifact-free EEG epochs used in the experiments had been manually selected by two skilled neurophysiologists. In practice, a fully automated diagnostic system is needed, thus automated artifact removal techniques, such as independent component analysis [28], need to be explored; this is the focus of our ongoing work.

5. CONCLUSION

This paper has presented two updates to a recently-proposed spectro-temporal EEG-based biomarker for Alzheimer's disease detection. First, a signal-to-noise ratio improvement stage is implemented, which averages amplitude modulation features prior to classification, thus is akin to the epoch averaging done in EEG event related potential studies. Second, two variants of the amplitude modulation features are developed to account for per-EEG-band modulations. Experimental results have shown that optimal results are achieved once 15 epochs have been averaged; performance begins to degrade beyond this point. Significant gains in classification accuracy, sensitivity, and specificity can be achieved once SNR improvement and the newly proposed features are incorporated into the system. As an advantage over existing work, the proposed system allows for accurate early detection of mild AD, as well as reliable assessment of disease severity progression from mild to moderate. Once fully automated (i.e., once blind artifact removal algorithms are in place), the proposed system may be used to assist clinicians with early AD assessment, as well as open doors to future in-home remote rehabilitation and/or patient monitoring tools. This is particularly useful given the world's aging population and their desire to "age-in-place".

6. REFERENCES

[1] Alzheimer Association, “Alzheimer’s disease facts and figures: 2010 report,” *Alzheimers Dement*, vol. 6, no. 2, pp. 158–194, 2010.

[2] B. Leifer, “Early diagnosis of Alzheimer’s disease: clinical and economic benefits,” *J Am Geriatr Soc*, vol. 51, pp. 281–288, 2003.

[3] Terry, R. “Neuropathological changes in AD”. Chapter 29. *Progress in Brain Research*. (Ed) Elservier Science BV, Vol. 101, Svennerholm, AK, 1994.

[4] D. Parikh et al., “Ensemble-based data fusion for early diagnosis of Alzheimers disease,” *Proc Intl Conf IEEE-EMBS*, pp. 2479–2482, 2005.

[5] Sarazin, M., L.C. de Souza, S. Lehericy, and B. Dubois, “Clinical and research diagnostic criteria for Alzheimer's disease”. *Neuroimaging Clin N Am*., 22(1): p. 23-32,viii, 2012.

[6] Babiloni, C., et al., “Reactivity of cortical alpha rhythms to eye opening in mild cognitive impairment and Alzheimer's disease: an EEG study”. *J Alzheimers Dis*. 22(4): p. 1047-64, 2010.

[7] D Cibils, “Dementia and qEEG (Alzheimer’s disease)”. *Clin. Neurophysiol*. 54, 289–294, 2002.

[8] T Locatelli, M Cursi, D Liberati, M Franceschi, G Comi, “EEG coherence in Alzheimer’s disease”. *Electroenceph. Clin. Neurophysiol*. 106(3), 229–237, 1998.

[9] D Abasolo, R Hornero, C Gomez, M Garcia, M Lopez, “Analysis of EEG background activity in Alzheimer’s disease patients with Lempel-Ziv complexity and central tendency measure”. *Med. Eng. Phys*. 28(4), 315–322, 2006.

[10] M Buscema, P Rossini, C Babiloni, 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”. *Artif. Intell. Med*. 40(2), 127–141, 2007.

[11] P Rossini, M Buscema, M Capriotti, E Grossi, G Rodriguez, C Del Percio, C Babiloni, “Is it possible to automatically distinguish resting EEG data of normal elderly vs. mild cognitive impairment subjects with high degree of accuracy?” *Clin. Neurophysiol.* 119(7), 1534–1545, 2008.

[12] L Trambaiolli, T H Falk, F Fraga, R Anghinah, A Lorena, “EEG spectro-temporal modulation energy: a new feature for automated diagnosis of Alzheimer’s disease”, in *Proc Intl Conf IEEE-EMBC*, 1, Boston, USA, pp. 3828-3831, 2011.

[13] Tiago H Falk, Francisco J Fraga, Lucas Trambaiolli and Renato Anghinah. “EEG Amplitude Modulation Analysis for Semi-Automated Diagnosis of Alzheimer's Disease”, *EURASIP Journal on Advances in Signal Processing,* vol 192*,* 2012.

[14] Karl E. Misulis, Toufic Fakhoury. *Spehlmann's Evoked Potential Primer*. Butterworth-heinemann, 2001.

[15] Brucki, S., R. Nitrini, P. Caramelli, P. Bertolucci, and I. Okamoto. “Suggestions for utilization of the mini-mental state examination in Brazil”. *Arq Neuropsiquiatr*. 61((3B)): p. 777-781, 2003.

[16] McKhann G, D.D., Folstein M, Katzman R, Price D, Stadlan EM., “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*, 1984. 34(7): p. 939-44.

[17] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR*. 4th ed., Washington, DC, 2002.

[18] J Jeong, “EEG dynamics in patients with Alzheimer’s disease”. *Clin. Neurophysiol*. 115(7), 1490–1505, 2004.

[19] L Trambaiolli, A Lorena, F Fraga, P Kanda, R Anghinah, R Nitrini. “Improving Alzheimer’s disease diagnosis with machine learning techniques”. *Clinic. EEG Neurosci*. 42(3), 160–165, 2011.

[20] P Nunez, R Srinivasan, *Electric fields of the brain: the neurophysics of EEG*, Oxford University Press, USA, 2006.

[21] S Sanei, J Chambers, *EEG Signal Processing*, Wiley-Interscience, New York, 2007.

[22] B Boashash, *Time Frequency Signal Analysis and Processing: A comprehensive Reference*, Elsevier, Amsterdam, 2003.

[23] 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, Article ID 761891, 6 pages, 2011.

[24] C Lehmann, T Koenig, V Jelic, L Prichep, R John, L Wahlund, Y Dodge, T Dierks, “Application and comparison of classification of Alzheimer’s disease in electrical brain activity” (EEG). *J. Neurosci. Meth.* 161, 342–350, 2007.

[25] N Cristianini, J Shawe-Taylor, *An Introduction to SVM and Other Kernel-Based Learning Methods*, Cambridge University Press, Cambridge, 2000.

[26] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten; “The WEKA Data Mining Software: An Update”; *SIGKDD Explorations*, Volume 11, Issue 1, 2009.

[27] Morris JC, Storandt M, Miller JP, et al. “Mild cognitive impairment represents early-stage OSAS or Sleep Apnea”. *Arch. Neurol.* 58 (3): 397–405, 2001.

[28] C Melissant, A Ypma, E Frietman, C Stam, “A method for detection of Alzheimer’s disease using ICA-enhanced EEG measurements”. *Artif. Intell. Med.* 33(3), 209–222, 2005.