

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/51598617>

# Improving Alzheimer's Disease Diagnosis with Machine Learning Techniques

**Article** in *Clinical EEG and neuroscience: official journal of the EEG and Clinical Neuroscience Society (ENCS)* · July 2011

DOI: 10.1177/155005941104200304 · Source: PubMed

CITATIONS

35

READS

741

6 authors, including:



**Lucas Trambaiolli**

Harvard Medical School

28 PUBLICATIONS 149 CITATIONS

[SEE PROFILE](#)



**Ana Carolina Lorena**

Instituto Tecnológico de Aeronautica

109 PUBLICATIONS 1,028 CITATIONS

[SEE PROFILE](#)



**Francisco J. Fraga**

Universidade Federal do ABC (UFABC)

50 PUBLICATIONS 249 CITATIONS

[SEE PROFILE](#)



**Paulo Medeiros Kanda**

Neurovale

27 PUBLICATIONS 203 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



An electrophysiological biomarker of memory for Alzheimer's disease [View project](#)



Functionality and cognitive performance of patients with behavioral variant Frontotemporal Dementia [View project](#)

# Clinical EEG and Neuroscience

<http://eeg.sagepub.com/>

---

## Improving Alzheimer's Disease Diagnosis with Machine Learning Techniques

Lucas R. Trambaiolli, Ana C. Lorena, Francisco J. Fraga, Paulo A.M. Kanda, Renato Anghinah and Ricardo Nitrini

*Clin EEG Neurosci* 2011 42: 160

DOI: 10.1177/155005941104200304

The online version of this article can be found at:

<http://eeg.sagepub.com/content/42/3/160>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



[EEG and Clinical Neuroscience Society](#)

**Additional services and information for *Clinical EEG and Neuroscience* can be found at:**

**Email Alerts:** <http://eeg.sagepub.com/cgi/alerts>

**Subscriptions:** <http://eeg.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Jul 1, 2011

[What is This?](#)

# Improving Alzheimer's Disease Diagnosis with Machine Learning Techniques

Lucas R. Trambaiolli, Ana C. Lorena, Francisco J. Fraga, Paulo A.M. Kanda, Renato Anghinah and Ricardo Nitrini

## Key Words

Alzheimer's Disease  
Coherence  
Electroencephalogram  
Support Vector Machines

## ABSTRACT

There is not a specific test to diagnose Alzheimer's disease (AD). Its diagnosis should be based upon clinical history, neuropsychological and laboratory tests, neuroimaging and electroencephalography (EEG). Therefore, new approaches are necessary to enable earlier and more accurate diagnosis and to follow treatment results. In this study we used a Machine Learning (ML) technique, named Support Vector Machine (SVM), to search patterns in EEG epochs to differentiate AD patients from controls. As a result, we developed a quantitative EEG (qEEG) processing method for automatic differentiation of patients with AD from normal individuals, as a complement to the diagnosis of probable dementia. We studied EEGs from 19 normal subjects (14 females/5 males, mean age 71.6 years) and 16 probable mild to moderate symptoms AD patients (14 females/2 males, mean age 73.4 years). The results obtained from analysis of EEG epochs were accuracy 79.9% and sensitivity 83.2%. The analysis considering the diagnosis of each individual patient reached 87.0% accuracy and 91.7% sensitivity.

## INTRODUCTION

Alzheimer's disease (AD) is considered the main cause of dementia in Western countries.<sup>1</sup> It is characterized by memory loss and impairment of at least one cognitive function (calculation, praxis, gnosis, executive functions, language, etc.). Notwithstanding, there is not a specific test to define AD and definitive diagnosis can only be established on autopsy or biopsy.<sup>2</sup>

Therefore, AD diagnosis should be based upon clinical history, laboratory tests, neuroimaging, neuropsychological batteries and EEG. Ergo, new approaches are necessary to enable earlier and more accurate diagnosis and to follow treatment results. Currently, neuropsychological screenings have AD diagnostic accuracy ranging from 85 to 93% in university hospitals. Unfortunately, these cognitive batteries require experienced people and lengthy sessions.<sup>3</sup> Consequently, it is still necessary to have a biological marker to help in the early diagnosis of AD.

QEEG is a procedure in current clinical use. It is non-invasive, safe and offers a superior temporal resolution compared to fMRI, SPECT and PET. Therefore, QEEG may potentially be used as a tool to screen a large number of people for risk of AD.<sup>4,5</sup> The most usual findings in EEG visual analysis<sup>6</sup> of AD are the displacement of background frequency into delta and theta ranges and the decrease of alpha central frequency.<sup>7</sup> Accordingly, Sandmann<sup>8</sup> observed a direct correlation between the degree of cognitive impairment and the power of low frequency electrical activity in the qEEG.

## Spectral Analysis (SpecA) and AD

SpecA has been considered 71% to 81% sensitive to changes in AD EEG background.<sup>9-12</sup> Saletu et al.<sup>13</sup> found a localized temporal decrease of alpha and beta activities in AD and a more widespread distribution of slow cerebral rhythms in vascular dementia (VaD). Pucci et al.<sup>14</sup> proposed that the "alpha" rhythm could be a diagnostic AD marker, since there is a decrease in the alpha frequency to 6.0-8.0 Hz in mild AD patients.

## Coherence (Coh) and qEEG

Coh quantifies the covariance between pairs of signals (EEG electrodes) as a function of frequency<sup>16</sup> and it is a well-established method to quantify connectivity through the corpus callosum,<sup>16</sup> high Coh is related to the structural and functional integrity of the intra- and inter-hemispheric cortical connections.<sup>17</sup> Consequently, Coh is becoming widely used in AD studies. For example, Besthorn et al.<sup>18</sup> found central and frontal Coh decrease in theta, alpha and beta bands of AD patients. Studies from Leuchter et al.<sup>19</sup>, and Locatelli et al.<sup>20</sup> had similar results with decreased Coh in AD EEGs.

SpecA and Coh generate such a large amount of data that they are not suitable for visual analysis and comparisons. Therefore, Machine Learning (ML) techniques are among the new methods used to handle high dimensional datasets and differentiate normal individuals from AD patients.<sup>3,21-24</sup> Earlier diagnosis of AD is necessary to delay disease progression. Consequently, the development of new EEG diagnostic tools would be very helpful.

The objective of this study is to use a ML technique known as Support Vector Machines (SVM)<sup>25</sup> to develop models able to extract and classify digital EEG signal (dEEG) dataset patterns from probands previously diagnosed as controls or AD patients. This technique is known by its good generalization ability and robustness to process high dimensional data as EEG signals.<sup>26-29</sup>

## MATERIALS AND METHODS

### Subjects

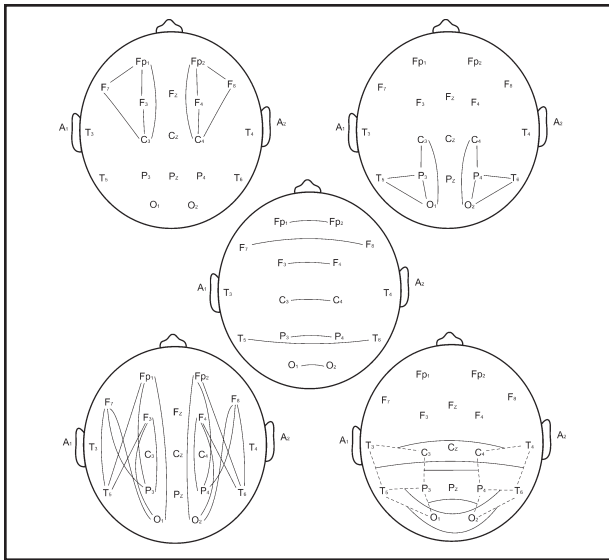
The dataset used in this study is composed of EEG signals (EEGs) recorded from two groups 60-80 years age-matched: (S1) 19 normal subjects, 14 females and 5 males, with averaged age of 71.6 years (sd 8.85); (S2) 16 probable AD patients (NINCDS-ADRDA criteria),<sup>30</sup> 14 females and 2 males, with averaged age of 73.4 (sd 6.12) years and mild to moderate symptoms (DSM-IV-TR).<sup>31</sup>

From the Mathematics, Computing and Cognition Center (CMCC) (L.R. Trambaiolli, A.C. Lorena), Engineering, Modeling and Applied Social Sciences Center (CECS) (F.J. Fraga), Universidade Federal do ABC (UFABC), São Paulo, Brazil; and the Reference Center of Behavioral Disturbances and Dementia (CEREDIC) of Medicine School (P. Kanda), University of São Paulo, São Paulo, Brazil.

Address correspondence and requests for reprints to Paulo A.M. Kanda, Reference Center of Behavioral Disturbances and Dementia (CEREDIC), Medicine School, University of São Paulo, Rua Arruda Alvim, 206 - CEP: 05.410-020 - São Paulo - SP - Brasil.

Email: pkanda@usp.br

Received: November 23, 2010; accepted: January 26, 2011.



**Figure 1.**

From left to right and from top to bottom: anterior coherences, posterior coherences, homolog coherences, long distance coherences, bipolar posterior coherences.

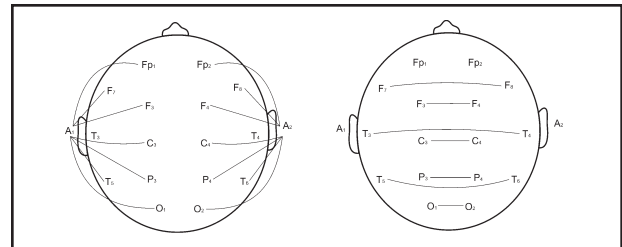
Patients and controls were submitted to the Brazilian version of the Mini-Mental State Examination (MMSE).<sup>32-33</sup> AD patients should score below 26 points in the test. Subjects had no history of diabetes mellitus, kidney disease, thyroid disease, alcoholism, liver disease, lung disease or vitamin B12 deficiency to exclude other causes of cognitive impairment.

#### Data acquisition and processing

The EEGs were recorded with a resolution of 12 bits, band pass of 1-30 Hz and sampling rate of 200 Hz. The acquisition equipment was the Braintech 3.0 (EMSA "Equipamentos Médicos") with impedance below 10 K. Electrodes were placed according to the International 10-20 System.<sup>5,34</sup> The inter-connected ear lobe electrodes reference used in this study is standard in our laboratory, despite the fact that there are controversies regarding which reference is the best.<sup>35-36</sup> Bipolar references were obtained virtually, through subtraction of the corresponding ear-lobe referenced (and digitized) signals. EEGs were recorded during 20 minutes with probands awake, relaxed with closed eyes. Two skilled neurophysiologists selected 8 seconds EEG fragments without artifacts (blinking, drowsiness, muscle movements or equipment-related artifacts) by visual inspection. This duration is appropriate because it provides good averaging of the parameters obtained with SpecA, as will be explained further.

A 512-point Fast Fourier Transform (FFT) was used for frequency analysis, applied after sub-segmenting the 8 sec epochs with Hamming windows of 2.5 s and 90% of overlap between successive windows,<sup>37</sup> thus generating six overlapped windows for each 8 sec EEG fragment. This window size was chosen to improve frequency resolution and it allows temporal resolution as well, since a sliding window (with 90% overlap) was also used in the signals. In order to eliminate the interference of the power grid (60Hz), EEG signals were filtered using an infinite impulse response low-pass elliptic filter with a cutoff frequency at 50Hz and a zero in the frequency of 60Hz.

The frequency bands were divided into sub-bands  $\delta 1(0,1-2,0\text{Hz})$ ,  $\delta 2(2,5-4,0\text{Hz})$ ,  $\theta 1(4,5-6,0\text{Hz})$ ,  $\theta 2(6,5-7,5\text{Hz})$ ,  $\alpha 1(8,0-10,0\text{Hz})$ ,  $\alpha 2(10,5-$



**Figure 2.**

From left to right: biauricular and bipolar electrode combinations to spectral peaks calculation.

12,0Hz),  $\beta 1(12,5-15\text{Hz})$ ,  $\beta 2(15,5-21,0\text{Hz})$ ,  $\beta 3(>21,0\text{Hz})$ .<sup>38</sup> The coherence between a pair of EEG channels for each epoch was obtained dividing the estimated cross spectrum of two channels by the auto-spectra of each channel<sup>39</sup> according to equation (1), where  $C_{ij}(\omega)$  is the cross spectral density and  $C_{ii}(\omega)$  and  $C_{jj}(\omega)$  are the power spectral densities of signals  $i$  and  $j$  (EEG channels). The average spectral windows of each epoch (using six overlapped windows) were calculated using the periodogram method of Welch.<sup>40</sup>

$$\text{Coh}_{ij}^2 = \frac{E | C_{ij}(\omega) |^2}{E | C_{ii}(\omega) | E | C_{jj}(\omega) |} \quad (1)$$

The pairs of electrodes used to estimate the cross spectrum coherence (Figure 1) were as follows: Inter-hemispheric<sup>38</sup> Fp1-Fp2, F7-F8, F3-F4, C3-C4, P3-P4, T5-T6 e O1-O2; Intra-hemispheric frontal electrodes<sup>20</sup>: Fp1-F7, Fp2-F8, Fp1-F3, Fp2-F4, Fp1-C3, Fp2-C4, F7-C3, F8-C4, F3-C3 e F4-C4; Intra-hemispheric rear electrodes<sup>20</sup>: O1-P3, O2-P4, O1-T5, O2-T6, O1-C3, O2-C4, P3-C3, P4-C4, T5-C3 e T6-C4; Equidistant electrodes<sup>20</sup>: O1-Fp1, O2-Fp2, O1-F7, O2-F8, O1-F3, O2-F4, P3-Fp1, P4-Fp2, P3-F7, P4-F8, P3-F3, P4-F4, T5-Fp1, T6-Fp2, T5-F7, T6-F8, T5-F3 e T6-F4; Rear electrodes in bipolar montage<sup>41</sup>: T3.C3-T4.C4, C3.P3-C4.P4, T5.P3-T6.P4, T3.T5-T4.T6, P3.O1-P4.O2, T5.O1-T6.O2.

In all combinations of electrode pairs the coherence operation is represented by a dash character (-) and the bipolar montage is represented by a dot character (.).

Another attribute used was the spectral peak, obtained from fast Fourier transform (FFT) spectral analysis.<sup>42</sup> The spectral peak (SPk) of each signal window can be defined as the point in EEG power spectral density (PSD) where spectral energy reaches its maximum value. Again, we used the six overlapped Hamming windows (2.5 seconds) to obtain the average spectral peaks of each epoch (one peak per frequency band). SPk was calculated with the following electrodes (Figure 2): Biauricular reference: Fp1, Fp2, F3, F4, F7, F8, C3, C4, P3, P4, T5, T6, O1 and O2; Bipolar reference: F3.F4, F7.F8, C3.C4, T3.T4, P3.P4, T5.T6 and O1.O2.

#### Support Vector Machines (SVMs)

SVMs constitute a Machine Learning (ML) technique based on the Statistical Learning Theory.<sup>43</sup> SVMs separates data by a hyperplane, considering bounds in the generalization ability of a linear classifier.<sup>44</sup> Accordingly, given a training data set  $T$  containing  $n$  pairs  $(\mathbf{x}_i, y_i)$ , where  $\mathbf{x}_i \in \mathbb{R}^m$ , is a data point with  $m$  dimensions (for instance, the features extracted from the EEG exam of a given patient) and  $y_i \in \{-1; +1\}$  is the class of  $\mathbf{x}_i$  (the diagnosis: -1 for normal and +1 for AD), SVMs seek the linear classifier  $g(\mathbf{x}) = \text{sgn}(\omega \cdot \mathbf{x} + b)$  separating data from classes +1 and -1 with minimum error while also maximizing the margin of separation between these classes (Figure 3).<sup>25</sup> In this paper the characteristics extracted from each EEG represent data points in two separated classes

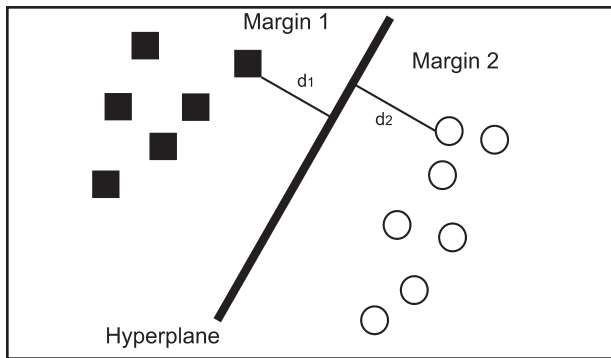


Figure 3.

Example of a linear frontier induced by a SVM. Each data point represents an EEG of a patient, while  $d_1$  and  $d_2$  represent the margins of separations between the classes according to the presence or absence of AD.

according to the presence or absence of AD. The maximization of the margin is accomplished through the minimization of the norm  $\|\omega\|$ . The following optimization problem is solved in this process:

$$\text{Minimize: } \|\omega\|^2 + C \sum_{i=1}^n \xi_i \quad (2)$$

Under the restrictions:  $y_i(\omega x_i + b) \geq 1 - \xi_i$  and  $\xi_i \geq 0$ , for  $i=1, \dots, n$  (3)

where  $C$  is a constant that imposes a different weight for training error over the generalization of the classifier and  $\xi_i$  are slack variables. The restrictions are originally imposed to ensure that no training data should be within the margins of separation between the classes. The slack variables relax these restrictions on the margins, in order to avoid an overfitting to training data and also for dealing with noisy data. The number of training errors and data between the margins is controlled by the minimization of the summation term in (2).<sup>25</sup>

$$\sum_{i=1}^n \xi_i$$

Nevertheless, the classifier obtained is still limited, since there are many data sets where data cannot be satisfactorily divided by a hyperplane, making a non-linear frontier more adequate to the problem.<sup>25</sup> The Cover theorem states that, if it is possible to increase the dimensions of data through a non-linear mapping function  $\Phi$  with a high probability they will become linearly separable in the new space, which is usually called feature space.<sup>45</sup>

The mathematical tool employed for the computation of  $\Phi$  is named Kernel. The Kernel  $K$  is a function which takes two variables  $x_i$  and  $x_j$ , representing two data points, and calculates the dot product between them in the feature space. Since all computations involving data points in SVMs are in the form of dot products, the non-linearization of SVMs can be easily accomplished through the use of a proper Kernel function. In this paper the non-linear RBF (Radial-Basis Function) Kernel function was used:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (4)$$

#### SVMs, how does it work?

We can illustrate the use of SVMs with a red-and-blue candies model. Suppose that you want to classify candies in two different classes (blue or red) using other characteristics than colors (attributes). It is assumed that each different class (color) has somewhat specific attributes (different colors have: 1. different concentrations of chocolate and peanut butter; 2. different shapes, round or square). These attributes represent the vector  $x_i$ .

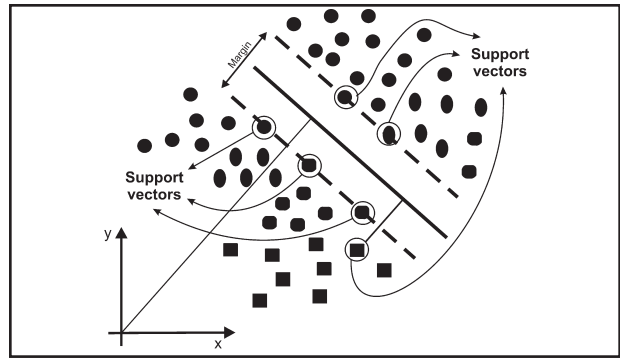


Figure 4.

Induced SVM in the red-and-blue candies example. The red candies closest to blue candies and the blue candies closest to the red ones are designed as support vectors used to establish the margin between the two distinct groups.

First, a candy bag is used as a SVM training set. The candies are placed on a table accordingly with their attributes (as in a Cartesian coordinate system). They are horizontally-distributed ( $x$  axis, Figure 4). Those with more peanut butter are selected to stay on the right and those with more concentration of chocolate are put on the left. Square candies are put on the top, round candies stay below ( $y$  axis, Figure 4).

After this classification colors are revealed. Suppose that blue and red candies are clearly separated. Some of these candies are classified as Support Vectors (SVs) when they skirt the border between the two colors. Consequently, it is possible to trace a line equidistantly to blue and red SVs in order to minimize separation error. These line coordinates define the training set separation model. As a result, you have separated blue and red candies in distinct classes based upon attributes other than the colors (Figure 4). Finally, the candies training set is removed from the table and the predefined coordinates are used to classify candies from other bags. A boundary line between different data classes is present in all machine learning techniques. The main characteristic of the linear SVM is to trace a maximum boundary linear hyperplane.<sup>25</sup>

#### SVMs induction

All SVMs in this paper were induced with the Weka tool<sup>45</sup> using default parameters values to allow fair comparisons among different features combinations. The parameters employed were  $C = 1,0$  and  $\gamma = 0,01$  in the RBF Kernel.

The 3305 EEG epochs dataset were divided in two parts, one to classification model training (68.06%) and other to test the induced model (31.94%). The experiments were repeated three times, with different data separations, in order to verify the variability of the SVM results to distinct data partitions. Patient data used for training (learning phase) was not used for testing and vice-versa.

#### RESULTS

Several different combinations of features (characteristics) extracted from the EEG exams were considered as input for the SVMs: frontal coherences, rear coherences, long distances coherences, homologous coherences, rear-bipolar coherences, biauricular peaks and all coherences combined to the biauricular peaks, bipolar peaks and all coherences combined to the bipolar peaks (Table 1). We measured the mean accuracy rates (percentage of epochs correctly classified) of the models obtained for each feature combination in the test sets, as well as their AUC (Area Under the ROC Curve performance measure), sensitivity and specificity in the classifications. Standard

**Table 1**

The mean and standard deviation performance measures for different partitions of data in the epochs classification of the EEG exams: accuracy, AUC (Area Under the ROC Curve), sensitivity (correct classifications for epochs from patients with AD) and specificity (correct classifications for epochs from normal patients)

Data	Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)
Frontal coherences	65,2 ± 3,5	0,66 ± 0,01	69,1 ± 9,4	63,2 ± 12,0
Posterior coherences	67,4 ± 3,4	0,68 ± 0,03	64,6 ± 3,5	70,4 ± 12,0
Long distances coherences	54,8 ± 2,4	0,55 ± 0,03	47,8 ± 8,5	62,3 ± 12,5
Homologous coherences	58,4 ± 9,2	0,58 ± 0,09	46,9 ± 21,5	69,1 ± 6,1
Posterior-bipolar	53,4 ± 8,0	0,52 ± 0,05	21,9 ± 16,8	<b>81,7 ± 12,7</b>
All coherences	72,4 ± 9,3	0,73 ± 0,09	77,1 ± 6,3	69,0 ± 12,1
Biauricular peaks	65,5 ± 8,6	0,65 ± 0,07	64,2 ± 12,1	66,5 ± 17,1
All coherences + biauricular peaks	72,4 ± 5,9	0,73 ± 0,04	77,0 ± 2,7	68,9 ± 11,1
Bipolar peaks	76,4 ± 3,6	0,76 ± 0,04	71,0 ± 0,1	80,5 ± 3,6
All coherences + bipolar peaks	<b>79,9 ± 3,9</b>	<b>0,80 ± 0,04</b>	<b>83,2 ± 3,6</b>	76,4 ± 8,5

Each line of the table represents a different data set configuration (different features extracted from the EEG signals).

The best results are boldfaced, while the worst results are in italics.

**Table 2**

For proband test diagnosis we used the ratio between the number of correctly classified epochs by the total number of epochs, by patient

Data	Accuracy (%)	Sensitivity (%)	Specificity (%)
All coherences	70,10 ± 8,71	69,17 ± 15,88	75,38 ± 6,88
Biauricular peaks	68,13 ± 8,49	60,83 ± 20,05	74,60 ± 9,91
All coherences + biauricular peaks	72,58 ± 14,17	75,83 ± 14,22	69,84 ± 14,55
Bipolar peaks	<b>86,97 ± 3,80</b>	<b>91,67 ± 14,43</b>	<b>84,92 ± 14,36</b>
All coherences + bipolar peaks	81,16 ± 5,86	82,50 ± 4,83	80,15 ± 7,65

Accuracy is the percentage of correct classification rate in the whole dataset. Sensitivity is probands correctly identified as AD. Specificity is probands correctly identified as normals.

deviation values of the results are also shown. Specificity corresponds to the ability to correctly classify AD epochs, while sensitivity regards the ability to correctly classify normal epochs. Higher sensitivity values can be considered better, since it prevents an AD patient from stopping treatment because of a wrong diagnosis.

The classifiers with better accuracies for the epochs (Table 1) were subsequently used to obtain the diagnosis of each individual subject. This analysis verifies the patient diagnosis accuracy rate, rather than the classification made considering each epoch. The results for coherence show that frontal and posterior configurations have better accuracy as observed by Locatelli.<sup>20</sup> On the other hand, posterior bipolar coherences have a good specificity, despite their low accuracy rate, as discussed by Trambaiolli.<sup>41</sup> The set of bipolar peaks presents better results than those obtained for the coherences.

For subjects' diagnoses, we considered the ratio between their number of correctly classified epochs by their total number of epochs. The accuracy, sensitivity and specificity were calculated considering the threshold of 50%+1 to diagnosis (Table 2).

## DISCUSSION

In Table 1 the results for coherence show that frontal and posterior configurations have better accuracy as observed by Locatelli.<sup>20</sup> On the other hand, posterior bipolar coherences have a good specificity, despite their low accuracy rate, as discussed by Trambaiolli.<sup>40</sup> The set of bipolar peaks presents better AD diagnostic results than those obtained for the coherences. Therefore, it is possible that the choice of the reference can influence the results. If we use all coherences and bipolar peaks together as input for a SVM classifier, an improvement is

noted in accuracy, AUC and sensitivity. Nevertheless, all these results refer to EEG epochs classification alone, which is not helpful for patient diagnosis. Spectrum peaks and combinations of these peaks with coherences achieved the best performance in the classification of the epochs, consequently, probands analyses were carried out for these particular datasets.

Both sets of features had accuracy improvement in the per proband evaluation scenario. For both datasets the accuracy, sensitivity and specificity rates were all above the level of 80%, and the sensitivity for the bipolar peaks data exceeds 90% (Table 2). The alpha and beta bands slowing of AD patients' EEG,<sup>47-48</sup> become evident when the difference between more localized electrical potentials is considered, i. e., between electrodes next to each other, as in bipolar montages. Despite the fact that sensitivity and specificity have high standard deviation values (sdv), a sensitivity of 91.7% is important, since it can reduce the risk of AD patients stopping treatment by an incorrect diagnosis.

Table 1 shows, for posterior-bipolar coherence tests, differences between values of specificity and sensitivity in favor of normal patients in detriment of AD patients. In the case of all coherences + bipolar peaks, there is a balance between sensitivity and specificity, suggesting that no class is strongly favored in detriment of the other.

It is also noteworthy that the sdv values obtained in the per patient analysis are quite high. Nevertheless, this can be attributed to the fact that we have relatively few patients for testing (13±2), and consequently, an elevation in sensitivity and specificity sdv. Therefore, the misclassification of only one patient results in a difference of total performance of approximately 15%.



## CONCLUSION

Although more tests are needed, involving a larger number of subjects, this study demonstrated that multiple combination of coherences and spectral peaks are good input features for SVM classifiers for automate AD diagnosis. We propose the use of SVM in the EEG study of demented patients because of their strong ability of generalization and robustness to work with high dimensional data, as these represented. The results indicate that the best feature inputs are: the combination of frontal coherences, rear coherences and long distances coherences<sup>20</sup>; homologous inter-hemispheric coherences<sup>38</sup>; posterior bipolar coherences, Trambaiolli<sup>41</sup>; and the peaks of spectrum.

Nevertheless, the main result in this paper is that the classification per patient had best results when we used as inputs for SVMs only the

frequencies of the peaks of spectrum obtained by bipolar recording. Therefore, we suggest that the use of bipolar peaks is a good tool providing a set of characteristics from EEG signals for SVMs in the classification of patients with AD. Given the simplicity of this pre-processing, allied to the high sensitivity obtained in the classification experiments, this can be considered a promising result.

Future work shall consider tuning the parameter values of the SVM classifiers, since this procedure can lead to a further increase in the classification performances achieved.

## DISCLOSURE AND CONFLICT OF INTEREST

L.R. Trambaiolli, A.C. Lorena, F.J. Fraga, P.A.M. Kanda, R. Anghinah and R. Nitrini have no conflicts of interest in relation to this article.

## REFERENCES

- Bird TD. Alzheimer's disease and other primary dementias. In: Braunwald E, Fauci AS, Kasper DL, Hauser SL, Longo DL, Jameson JL. (eds). *Harrison's Principles of Internal Medicine*. New York: McGraw-Hill; 2001: 2391-2399.
- Terry RD. Neuropathological changes in Alzheimer's disease. *Prog Brain Res* 1994; 101: 383-390.
- Parikh D, Stepenosky N, Topalis A, Green D, Kounios J, Clark C, Polikar R. Ensemble based data fusion for early diagnosis of Alzheimer's disease. *P Ann Int IEEE EMBS* 2005; 2479-2482.
- Luccas FLC, Braga NIO, Fonseca LC, Frochtengarter ML. Recomendações para o registro e interpretação do mapeamento topográfico do eletroencefalograma (EEG) e potenciais evocados sensoriais (PES) parte I: aspectos gerais [in portuguese]. *J Epilepsy Clin Neurophysiol* 1996; 2: 175-182.
- Luccas FJ, Anghinah R, Braga NI, Fonseca LC, Frochtengarten ML, Jorge MS, Kanda PAM. Guidelines for recording/analyzing quantitative EEG and evoked potentials. Part II: Clinical aspects. *Arq Neuro-psiquiatr* 1999; 57(1): 132-146.
- Nuwer MR, Comi G, Emerson R, Fuglsang-Frederiksen J, Guérit M, Hinrichs H, et al. IFCN standards for digital recording of clinical EEG. *Electroencephalogr Clin Neurophysiol* 1998; 106: 259-261.
- Klass DW, Brenner RP. Electroencephalography of the elderly. *J Clin Neurophysiol* 1995; 12: 116-131.
- Sandmann MC, Piana ER, Sousa DS, Bittencourt PR. Digital EEG with brain mapping in Alzheimer's dementia and Parkinson's disease: a prospective controlled study. *Arq Neuro-psiquiatr* 1996; 54(1): 50-56.
- Dierks T, Perisic I, Frölich L, Ihl R, Maurer K. Topography of the qEEG in dementia of Alzheimer type: relation to severity of dementia. *Psychol Res-Psych FO* 1991; 40: 181-194.
- Martin-Loeches M, Gil P, Jimenez F, Exposito FJ, Miguel F, Cacabelos R, Rubis FJ. Topographic maps of brain electrical activity in primary degenerative dementia of Alzheimer type and multiinfarct dementia. *Biol Psychiat* 1991; 29: 211-223.
- Leuchter AF, Cook IA, Newton TF, Dunkin J, Walter DO, Rosenberg-Thompson S, et al. Regional differences in brain electrical activity in dementia: use of spectral power and spectral ratio measures. *Electroencephalogr Clin Neurophysiol* 1993; 87: 385-393.
- Anderer P, Saletu B, Klöppel B, Semlitsch HV, Werner H. Discrimination between demented patients and normals based on topographic EEG slow wave activity: comparison between z statistics, discriminant analysis and artificial neural network classifiers. *Electroencephalogr Clin Neurophysiol* 1994; 91: 108-117.
- Saletu B, Paulus E, Grunbergerer J. Correlation maps: on the relation of electroencephalographic slow wave activity to computerized tomography and psychopathometric measurements in dementia. In: Maurer K, (ed). *Imaging of Brain in Psychiatry and Related Fields*. Berlin: Springer-Verlag; 1993: 263-265.
- Pucci E, Belardinelli N, Cacchiò G, Signorino M, Angeleri F. EEG power spectrum differences in early and late onset forms of Alzheimer's disease. *Clin Neurophysiol* 1999; 110: 621-631.
- Shaw JC. An introduction to the coherence function and its use in EEG signal analysis. *J Med Eng Technol* 1981; 5(6): 279-288.
- Nielsen T, Montplaisir J, Lassonde M. Decreased interhemispheric EEG coherence during sleep in agenesis of the corpus callosum. *Eur Neurol* 1993; 33: 173-176.
- Claus JJ, Strijers RLM, Jonkman EJ, Ongerboer de Visser BW, Jonker C, Walstra GJM, et al. The diagnostic value of EEG in mild senile Alzheimer's disease. *Clin Neurophysiol* 1999; 110: 825-832.
- Besthorn C, Zerfass R, Geiger-Kabisch C, Sattel H, Daniel S, Schreiter-Gasser U, Forstl H. Discrimination of AD and normal aging by EEG data. *Electroencephalogr Clin Neurophysiol* 1997; 103: 241-248.
- Leuchter AF, Spar JE, Walter DO, Weiner H. Electroencephalographic spectra and coherence in the diagnosis of Alzheimer's-type and multi-infarct dementia. *Arch Gen Psychiat* 1987; 44: 993-998.
- Locatelli T, Cursi M, Liberati D, Franceschi M, Comi G. EEG coherence in Alzheimer's disease. *Electroencephalogr Clin Neurophysiol* 1998; 106: 229-237.
- Schettin V. Polynomial neural networks learn to classify EEG signals. *P NIMIA-SC2001* 2001; 9-20.
- Yagneswaran S, Baker M, Petrosian A. 2002. Power frequency and wavelet characteristics in differentiating between normal and Alzheimer EEG. *P Ann Int IEEE EMBS* 2002; 46-47.
- Cho SY, Kim BY, Park EH, Kim JE, Whang WW, Han SK, Kim HY. Automatic recognition of Alzheimer's disease with single channel EEG recording. *P Ann Int IEEE EMBS* 2003; 2655-2658.
- Kim TH, Kim BY, Park EH, Kim JW, Hwang EW, Han SK, Cho S. Computerized recognition of Alzheimer disease: EEG using genetic algorithms and neural network. *Future Gener Comp Sy* 2005; 21: 1124-1130.
- Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge: Cambridge University Press; 2000.
- Cichoki A, Shishkin SL, Musha T, Leonowicz Z, Asada T, Kurachi T. EEG filtering based on blind source separation (BSS) for early detection of Alzheimer's disease. *Clin Neurophysiol* 2005; 116: 729-737.

27. Stepenosky N, Topalis A, Syed H, Green D, Kounios J, Clark C, Polikar R. Boosting based classification of event related potentials for early diagnosis of Alzheimer's disease. *P Ann Int IEEE EMBS* 2005; 2494-2497.
28. Abe JM, Lopes HFS, Anghinah R. Paraconsistent artificial neural networks and Alzheimer's disease. *Dem Neuropsychol* 2007; 1:241-247.
29. Lehmann C, Koenig T, Jelic V, Prichet L, John RE, Wahlund LO, Dodge Y, Dierks T. Application and comparison of classification of Alzheimer's disease in electrical brain activity (EEG). *J Neurosci Meth* 2007; 161: 342-350.
30. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of AD: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on AD. *Neurol* 1984; 34: 939-944.
31. American Psychiatric Association. Diagnostic and Statistical Manual (DSM-IV-TR). 3rd ed.; 2000.
32. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *J Psychiat Res* 1975; 12: 189-198.
33. Brucki SM, Nitrini R, Caramelli P, Bertolucci PH, Okamoto IH. Suggestions for utilization of the mini-mental state examination in Brazil. *Arq Neuropsiquiatr* 2003; 61(3B): 777-781.
34. Jasper HH. The ten-twenty electrode system of the International Federation. *Electroencephalogr Clin Neurophysiol* 1958; 10: 371-373.
35. Rappelsberger P. EEG coherence and reference signals: experimental results and mathematical explanations. *Med Biol Eng Comput* 1998; 38(4): 399-406.
36. Yao D, Wang L, Arendt-Nielsen L, Chen AC. The effect of reference choices on the spatio-temporal analysis of brain evoked potentials: the use of infinite reference. *Comput Biol Med* 2007; 37(11): 1529-1538.
37. Sanei S, Chambers JA. EEG Signal Processing. New-York: Wiley-Interscience; 2007.
38. Anghinah R. Estudo da densidade espectral e da coerência do eletroencefalograma em indivíduos adultos normais e com doença de Alzheimer provável [in portuguese]. Ph.D. dissertation, São Paulo: Univ of São Paulo; 2003.
39. Gotman J, Gloor P, Ray WF. A quantitative comparison of traditional reading of the eeg and interpretation of computer-extracted features in patients with supratentorial brains lesions. *Electroencephalogr Clin Neurophysiol* 1975; 38: 623-639.
40. Welch PD. The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE T Audio Electroacoust* 1967; 15(2): 70-73.
41. Trambaioli LR, Lorena AC, Fraga FJ, Anghinah R. Uso de aprendizado de máquina no auxílio ao diagnóstico de Alzheimer [in portuguese]. *Rev Elet Inic Cientif* 2009; 3: 1-15.
42. Raicher I, Takahashi DY, Kanda PAM, Nitrini R, Anghinah R. qEEG spectral peak in Alzheimer's disease. *Dem Neuropsychol* 2008; 2(1): 9-12.
43. Vapnik VN. The Nature of Statistical Learning Theory. New York: Springer-Verlag; 1995.
44. Campbell C. An introduction to kernel methods. *Stud Fuzz Soft Comp* 2001; 66: 155-192.
45. Haykin S. Neural Networks A Comprehensive Foundation. New Jersey: Prentice-Hall; 1999.
46. Witten IA, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. San Francisco: Morgan Kaufmann; 2005.
47. Yener GG, Leuchter AF, Jenden D, Read SL, Cummings JL, Miller BL. Quantitative EEG in frontotemporal dementia. *Clin Neurophysiol* 1996; 27(2): 61-68.
48. Cibils D. Dementia and qEEG (Alzheimer's disease). *Clin Neurophysiol Sup* 2002; 54: 289-294.