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:
SAGE
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

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.1 It is characterized by memory loss and impairment of at least one cognitive function (calculation, praxis, gnosia, 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.2 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.3 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.4-6 The most usual findings in EEG visual analysis4,6 of AD are the displacement of background frequency into delta and theta ranges and the decrease of alpha central frequency.7 Accordingly, Sandmann8 observed a direct frequency into delta and theta ranges and the decrease of alpha 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.9,10 Saletu et al.11 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.12 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 frequency16 and it is a well-established method to quantify connectivity through the corpus callosum.16 High Coh is related to the structural and functional integrity of the intra- and inter-hemispheric cortical connections.17 Consequently, Coh is becoming widely used in AD studies. For example, Besthorn et al.18 found central and frontal Coh decrease in theta, alpha and beta bands of AD patients. Studies from Leuchter et al.19, and Locatelli et al.20 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.3,11-24 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)25 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.16-20

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),26 14 females and 2 males, with averaged age of 73.4 (sd 6.12) years and mild to moderate symptoms (DSM-IV-TR).27
Patients and controls were submitted to the Brazilian version of the Mini-Mental State Examination (MMSE). 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. 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. 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, 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 12,0Hz, β1(12.5-15Hz), β2(15.5-21,0Hz), β3(>21,0Hz). 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 according to equation (1), where $C_{ij}(\omega)$ is the cross spectral density and $C_{ii}(\omega)$ 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. The pairs of electrodes used to estimate the cross spectrum coherence (Figure 1) were as follows: Inter-hemispheric: Fp1-Fp2, F7-F8, F3-F4, C3-C4, P3-P4, T5-T6 eO1-O2; Intra-hemispheric frontal electrodes: Fp1-F7, Fp2-F8, Fp1-F3, Fp2-F4, Fp1-C3, Fp2-C4, F7-C3, F8-C4, F3-C3 eF4-C4; Intra-hemispheric rear electrodes: O1-P3, O2-P4, O1-T5, O2-T6, O1-C3, O2-C4, P3-C3, P4-C4, T5-C3 eT6-C4; Equidistant electrodes: O1-Fp1, O2-Fp2, O1-F7, O2-F8, O1-F3, O2-F4, Fp1-Fp2, P3-F7, F4-F8, P3-F3, P4-F4, Fp1-Fp2, T6-Fp2, T5-F7, T6-F8, T5-F3 eT6-F4; Rear electrodes in bipolar montage: 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. 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. SVMs separates data by a hyperplane, considering bounds in the generalization ability of a linear classifier. Accordingly, given a training data set $T$ containing $n$ pairs $(x_i, y_i)$, where $x_i \in \mathbb{R}^m$, 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 $x_i$ (the diagnosis: -1 for normal and +1 for AD). SVMs seek the linear classifier $g(x) = \text{sgn}(\omega^T x + b)$ separating data from classes $+1$ and $-1$ with minimum error while also maximizing the margin of separation between these classes (Figure 3). In this paper the characteristics extracted from each EEG represent data points in two separated classes.
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 \( \mathbf{x}_i \).

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.\(^{25}\)

### SVMs induction

All SVMs in this paper were induced with the Weka tool\(^{45}\) 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
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.20 On the other hand, posterior-bipolar coherences have a good specificity, despite their low accuracy rate, as discussed by Trambaioli.40 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.20 On the other hand, posterior-bipolar coherences have a good specificity, despite their low accuracy rate, as discussed by Trambaioli.40 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.

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.

**Table 1**

<table>
<thead>
<tr>
<th>Data</th>
<th>Accuracy (%)</th>
<th>AUC</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frontal coherences</td>
<td>65.2 ± 3.5</td>
<td>0.66 ± 0.01</td>
<td>69.1 ± 9.4</td>
<td>63.2 ± 12.0</td>
</tr>
<tr>
<td>Posterior coherences</td>
<td>67.4 ± 3.4</td>
<td>0.68 ± 0.03</td>
<td>64.6 ± 3.5</td>
<td>70.4 ± 12.0</td>
</tr>
<tr>
<td>Long distances coherences</td>
<td>54.8 ± 2.4</td>
<td>0.55 ± 0.03</td>
<td>47.8 ± 8.5</td>
<td>62.3 ± 12.5</td>
</tr>
<tr>
<td>Homologous coherences</td>
<td>58.4 ± 9.2</td>
<td>0.58 ± 0.09</td>
<td>46.9 ± 21.5</td>
<td>69.1 ± 6.1</td>
</tr>
<tr>
<td>Posterior-bipolar</td>
<td>53.4 ± 8.0</td>
<td>0.52 ± 0.05</td>
<td>21.9 ± 16.8</td>
<td>81.7 ± 12.7</td>
</tr>
<tr>
<td>All coherences</td>
<td>72.4 ± 9.3</td>
<td>0.73 ± 0.09</td>
<td>77.1 ± 6.3</td>
<td>69.0 ± 12.1</td>
</tr>
<tr>
<td>Biauricular peaks</td>
<td>65.5 ± 8.6</td>
<td>0.65 ± 0.07</td>
<td>64.2 ± 12.1</td>
<td>66.5 ± 17.1</td>
</tr>
<tr>
<td>All coherences + biauricular peaks</td>
<td>72.4 ± 5.9</td>
<td>0.73 ± 0.04</td>
<td>77.0 ± 2.7</td>
<td>68.9 ± 11.1</td>
</tr>
<tr>
<td>Bipolar peaks</td>
<td>76.4 ± 3.6</td>
<td>0.76 ± 0.04</td>
<td>71.0 ± 0.1</td>
<td>80.5 ± 3.6</td>
</tr>
<tr>
<td>All coherences + bipolar peaks</td>
<td>79.9 ± 3.9</td>
<td>0.80 ± 0.04</td>
<td>83.2 ± 3.6</td>
<td>76.4 ± 8.5</td>
</tr>
</tbody>
</table>

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**

<table>
<thead>
<tr>
<th>Data</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All coherences</td>
<td>70.10 ± 8.71</td>
<td>69.17 ± 15.88</td>
<td>75.38 ± 6.88</td>
</tr>
<tr>
<td>Biauricular peaks</td>
<td>68.13 ± 8.49</td>
<td>60.83 ± 20.05</td>
<td>74.80 ± 9.91</td>
</tr>
<tr>
<td>All coherences + biauricular peaks</td>
<td>72.58 ± 14.17</td>
<td>75.83 ± 14.22</td>
<td>69.84 ± 14.55</td>
</tr>
<tr>
<td>Bipolar peaks</td>
<td>86.97 ± 3.80</td>
<td>91.67 ± 14.43</td>
<td>84.92 ± 14.36</td>
</tr>
<tr>
<td>All coherences + bipolar peaks</td>
<td>81.16 ± 5.86</td>
<td>82.50 ± 4.83</td>
<td>80.15 ± 7.65</td>
</tr>
</tbody>
</table>

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.
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 automatic 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; homologous inter-hemispheric coherences; posterior bipolar coherences, Trambaiolli; 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. Trambaioli, A.C. Lorenna, F.J. Fraga, P.A.M. Kanda, R. Anghinah and R. Nitrini have no conflicts of interest in relation to this article.

REFERENCES


