SPECTRAL ANALYSIS OF PREDICTIVE ERROR IN ALZHEIMER'S DISEASE DIAGNOSTICS

L. Tylová[*], J. Kukal[*], O. Vyšata[†]

Abstract: The new method is based on non-linear one-step predictor, which is designed as MLP neural network. It is a kind of low-pass non-linear filter. The difference between raw EEG and the ANN output is then a subject of band spectral analysis. The differences in this power spectrum between Alzheimer's diseased and control patient group are statistically significant.

Key words: *EEG, Alzheimer's disease, ANN, model error, spectral analysis*

Received: October 24, 2010
Revised and accepted: November 2, 2013

1. Introduction

Alzheimer's disease (AD) is the most common dementia. This disease affects approximately 7 % of people older than 65 years and 40 % of people older than 80 years [7]. Dementia is characterized by memory decline and others neurophysiological changes that occur in the elderly and the risk of disease increases rapidly with age [7].

EEG signals reflect the bioelectrical activity of the brain. Electroencephalographic records are one of the tools for diagnosis of neurological diseases. Traditional analysis relies mainly on detection of spectral changes: performs the analysis of selected frequency bands, then calculate the corresponding spectral powers, whose changes may indicate dysfunction of the nervous system.

Analysis of the power spectrum of a healthy active brain suggests [2] that there are four main frequency bands: δ (0.5–4 Hz), θ (5–8 Hz), α (9–12 Hz), β (13–20 Hz). In the frequency domain, there are established following differences in the EEG records of healthy patients and patients with Alzheimer's dementia: an increase in theta and delta rhythms, decline in beta rhythm and slowing of alpha rhythm.

[*]L. Tylová, J. Kukal
Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague, tylovluc@fjfi.cvut.cz
[†]O. Vyšata
Dept. of Neurology, Faculty Hospital Hradec Kralove

Many works [3], [4], [5] also show the possibility of using artificial intelligence to solve the problem of identification of Alzheimer's disease. Among them those employing neural network to address this problem. Dominating amount of the works has many common features: perform artefacts cleaning of EEG record, perform the analysis of time series decomposition to frequency bands, perform further processing of these bands. After all these operations neural network is used for classification purposes.

The possibility of classification of healthy people and people with Alzheimer's disease, in this study, is explored using the following assumptions:

- don't use EEG signals adjusted with artefacts filtering;
- use neural network to detect patters in the EEG signal;
- prediction error is used for classification purposes, this prediction error was obtained from the use of neural network, trained on human health and subsequently used for a person with Alzheimer's disease;
- don't use an artificial neural network for classification purpose.

2. Alzheimer's Disease Diagnosis via EEG

Electroencephalography is a continuous multi-channel recording of electrical potential difference measured by electrodes placed on the scalp. EEG was first introduced and described by Berger [6]. Data which are collected from a typical EEG experiment are a sequence of time points sampled at 128–1024 Hz in general.

EEG recordings were obtained from 16 healthy people and 16 people with Alzheimer's disease. All patients were examined laying on the bed in the darkened room, they were at a quiet state and had their eyes closed. Measurements were performed using 19 active + 2 reference (ear) electrodes placed on the surface of the head according to the international 10/20 system. Sampling frequency was 200 Hz. Total number of patients is relatively small for detailed statistical analysis. But the data source of EEG series was limited to 32 samples. Therefore, there will be applied only standard t-test and ROC for data analysis.

3. ANN as Intelligent Filter of EEG

3.1 Signal Description

One of the main points of this work is to construct a model that would allow to assess the general patterns of EEG recording in healthy people and would create preconditions for the decision rules required to classify EEG recordings of healthy people and people with Alzheimer's disease.

Formulation of the problem is largely procedural in nature, taking into account only one factor, the one that EEG record reflects changes in brain bioelectrical activity, and among these changes are those that carry important information for us.

In the most general case, the behavior of EEG signal can be described as a superposition of a function z which describes important information for us, and some

random component e. The estimation $\hat{z}$ of the signal function z will be implemented by using neural network, in this paper. Random component $\hat{e}$, which arises as a result of this assessment, will serve as the calculation of the noise component e.

3.2 Neural Network

Let $n \in \mathbb{N}$ be number of inputs, $N \in \mathbb{N}$ be number of outputs and $H \in \mathbb{N}$ be number of neurons in the hidden layer. Let $\mathrm{x} \in \mathbb{R}^{\mathrm{n}}$ be input vector, $\mathrm{y} \in \mathbb{R}^{\mathrm{N}}$ be output vector and $\mathrm{h} \in \mathbb{R}^{\mathrm{H}}$ be signal vector in the hidden layer. The three layer ANN – multi-layer perceptron (MLP) operates according to equations

$$\mathrm{h} = \mathsf{f}(\mathbb{W}\mathrm{x} + \mathrm{w}_0) \tag{1}$$

$$\hat{\mathrm{z}} = \mathbb{V}\mathrm{h} + \mathrm{v}_0 \tag{2}$$

where $\mathbb{W} \in \mathbb{R}^{H \times n}$, $\mathbb{V} \in \mathbb{R}^{N \times H}$ are weight matrices, $\mathrm{w}_0 \in \mathbb{R}^{\mathrm{H}}$, $\mathrm{v}_0 \in \mathbb{R}^{\mathrm{N}}$ are biases and f is a non-polynomial function. After the decompositions:

$$\mathbb{W} = \begin{pmatrix} \mathrm{w}_1 \\ \vdots \\ \mathrm{w}_{\mathrm{H}} \end{pmatrix}, \mathbb{V} = \begin{pmatrix} \mathrm{v}_1 \\ \vdots \\ \mathrm{v}_{\mathrm{N}} \end{pmatrix}$$

We can establish the vector of ANN parameters

$$\mathrm{p} = (\mathrm{w}_0^{\mathrm{T}}, \mathrm{w}_1^{\mathrm{T}}, ..., \mathrm{w}_{\mathrm{H}}^{\mathrm{T}}, \mathrm{v}_0^{\mathrm{T}}, \mathrm{v}_1^{\mathrm{T}}, ..., \mathrm{v}_{\mathrm{N}}^{\mathrm{T}})$$

consisting of $M = (n+1)H + (H+1)N$ real coordinates. The resulting MLP as ANN can be formally rewritten as

$$\mathrm{y} = \mathsf{ANN}(\mathrm{x}, \mathrm{p}) \tag{3}$$

3.3 Learning Strategy

Let $m \in \mathbb{N}$ be number of patterns for ANN learning. The pattern set can be represented via matrices

$$\mathbb{X} = \begin{pmatrix} \mathrm{x}_1 \\ \vdots \\ \mathrm{x}_{\mathrm{m}} \end{pmatrix} \in \mathbb{R}^{m \times n}, \mathbb{Y} = \begin{pmatrix} \mathrm{y}_1 \\ \vdots \\ \mathrm{y}_{\mathrm{m}} \end{pmatrix} \in \mathbb{R}^{m \times N}$$

The method of least squares was used to ANN learning. Adequate objective function

$$\mathsf{F}(\mathrm{p}) = \sum_{\mathrm{k}=1}^{\mathrm{m}} \|\mathrm{y}_k - \mathsf{ANN}(\mathrm{x}_k, \mathrm{p})\|^2 \tag{4}$$

is thus subject of minimization.

Due to multi-modality of $\mathsf{F}(\mathrm{p})$ we applied Fast Simulated Annealing (FSA) method. The algorithm of FSA produces a parameter sequence p_0, p_1,..., $\mathrm{p}_k \in \mathbb{R}^M$ beginning with initial vector p_0 according to the rule:

$$\mathrm{p}_{k+1} = \mathrm{p}_k^* \quad \text{when} \quad \mathsf{F}(\mathrm{p}_k^*) < \mathsf{F}(\mathrm{p}_k) + T_k \tan \rho_k$$

$$\mathrm{p}_{k+1} = \mathrm{p}_k^* \quad \text{when} \quad \mathsf{F}(\mathrm{p}_k^*) \geq \mathsf{F}(\mathrm{p}_k) + T_k \tan \rho_k \tag{5}$$

where $\mathrm{p}_k^* = \mathrm{p}_k + gT_k \cdot \tan \mathrm{r}_k, \rho_k, \mathrm{r}_{k,j} \sim \mathrm{U}(-\frac{\pi}{2}, +\frac{\pi}{2})$ are independent uniformly distributed random variables, $g > 0$ is scaling factor and $T_k > 0$ is dimensionless temperature. The strategy of FSA cooling is

$$T_k = \frac{T_0}{1 + \lfloor \frac{k}{r} \rfloor} \tag{6}$$

where $r \in \mathbb{N}$ is a repeating period.

4. Spectral Analysis of EEG Signal

Power spectrum describes the energy distribution of the frequencies of the dynamic system. The dynamic development of simple systems can usually be described by a certain frequency range. An opposite situation is typical for complex systems: cannot be selected any particular frequency band. Frequency components were processed in each of the four frequency bands using the following relationship.

$$r_{band} = \frac{\sum\limits_{band} |\mathrm{fft}|^2}{\sum |\mathrm{fft}|^2} \cdot 100\,\% \tag{7}$$

where fft is a result of application of the fast Fourier transform to analyse EEG signal, *band* is one of the four main frequency bands: δ, θ, α or β. Due to the large non-linearity of EEG signal, Fourier transformation was applied not only to the data itself but also on the resulting prediction error obtained by using neural networks. It is natural to expect that the non-linearity error is smaller than the non-linearity of the original signal.

5. Results

To eliminate noise in the EEG data, EEG signals of all patients were analysed in the range of indices from 20000 to 50000. Identify patterns in the EEG recording was performed by one-step prediction using MLP neural network. The used MLP network consisted of one hidden layer with four hidden neurons. Hyperbolic tangent function was used as an activation function of MLP network. As a standard healthy person has been chosen one patient (pivot) whose EEG signal had the average statistical characteristics regarding the set of healthy people. According to [7] relative power in four frequency bands were calculated and then we obtained four descriptors per each healthy patient. Average patient feature was obtained by coordinate-wise averaging of relative power. Pivot healthy person was selected

using 1-NN method as the nearest one to the average healthy patient. The neural network was trained on the EEG signal of this patient, generalization abilities of used neural network were tested on the EEG signals of the remaining healthy patients. Subsequently, the neural network was applied to the EEG signals of patients with Alzheimer's disease. All electrodes were used for the prediction of EEG signals using MLP network in the corresponding EEG signals of pivot patient. Therefore, MLP ANN consists of 19 inputs (all channels in current step), 4 hidden neurons, and one output neuron (given channel in next step). Every channel was analyzed separately yielding 19 MLP ANN predictors. Training all neural network were performed by FSA algorithm, containing 300 interior and 300 exterior cycles. Classification of healthy patients and patients with Alzheimer's disease was based on a review of a one-step error prediction for EEG signal, using MLP neural network which was taught on EEG signal of a healthy pivot patient, and then used to predict EEG signal of other patients. In accordance with the above definitions, we have: $n = 19$, $H = 4$, $N = 1$, $T_0 = 0.001$, $g = 1$, $r = 300$, $k_{max} = 300$.

Results of band spectral analysis of individual channels for raw EEG data are collected in the Tab. I. Individual p-values are results of two-sample two-sided t-test of hypothesis H_0 that relative power (for given frequency band and channel) is the same for AD and CN group of patients. Adequate ROC diagram is depicted on the Fig. 1 for four bands and first channel. Tab. II shows results of the t-test of significance for each band, in the power spectrum of EEG signal, originating from the first channel, at a significance level of 5 %.

ch	δ	θ	α	β
1	9.4×10^{-2}	1.6×10^{-2}	1.1×10^{-1}	2.6×10^{-2}
2	3.2×10^{-1}	1.9×10^{-2}	2.4×10^{-2}	1.1×10^{-1}
3	3.6×10^{-1}	3.3×10^{-2}	$\mathbf{1.6 \times 10^{-4}}$	1.4×10^{-1}
4	1.4×10^{-1}	8.1×10^{-3}	2.4×10^{-3}	3.8×10^{-1}
5	2.1×10^{-1}	6.8×10^{-3}	1.7×10^{-2}	8.0×10^{-2}
6	3.7×10^{-1}	5.2×10^{-3}	7.0×10^{-3}	1.3×10^{-1}
7	8.8×10^{-1}	8.1×10^{-2}	3.9×10^{-3}	2.7×10^{-1}
8	5.2×10^{-1}	9.7×10^{-2}	$\mathbf{3.0 \times 10^{-5}}$	6.2×10^{-1}
9	7.8×10^{-1}	3.2×10^{-1}	$\mathbf{3.9 \times 10^{-4}}$	2.9×10^{-2}
10	9.8×10^{-2}	3.2×10^{-3}	$\mathbf{3.5 \times 10^{-4}}$	7.2×10^{-2}
11	3.7×10^{-1}	9.2×10^{-2}	$\mathbf{9.0 \times 10^{-4}}$	1.6×10^{-1}
12	5.8×10^{-1}	6.9×10^{-2}	$\mathbf{9.8 \times 10^{-4}}$	8.0×10^{-1}
13	6.5×10^{-1}	5.4×10^{-1}	$\mathbf{2.0 \times 10^{-4}}$	6.9×10^{-2}
14	8.4×10^{-1}	2.8×10^{-1}	$\mathbf{8.0 \times 10^{-5}}$	1.2×10^{-2}
15	3.0×10^{-1}	8.3×10^{-1}	$\mathbf{7.6 \times 10^{-4}}$	$\mathbf{7.2 \times 10^{-4}}$
16	2.5×10^{-1}	7.2×10^{-2}	$\mathbf{3.0 \times 10^{-5}}$	1.1×10^{-1}
17	5.5×10^{-1}	3.9×10^{-2}	2.0×10^{-3}	5.0×10^{-2}
18	7.7×10^{-1}	7.0×10^{-1}	9.0×10^{-3}	1.3×10^{-3}
19	7.8×10^{-1}	3.0×10^{-1}	$\mathbf{9.9 \times 10^{-4}}$	3.1×10^{-2}

Tab. I *P-values for significant differences (AD × CN) in the case of raw EEG (t-test).*

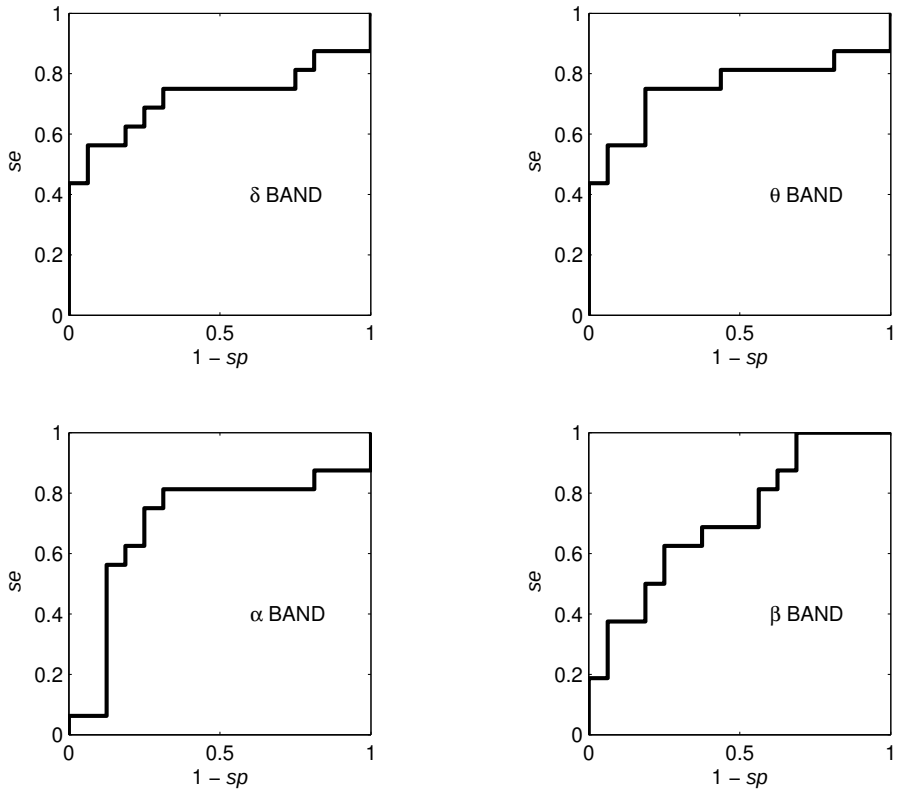

Fig. 1 *ROC for the 1st channel and $\delta, \theta, \alpha, \beta$ bands in the case of raw EEG data.*

The differences between raw EEG and the output of ANN is called here as prediction error. The best are the results of band spectral analysis of the prediction error. Relative power of prediction in given channel and band was also subject of statistical testing. Results are involved in the Tab. III. Adequate ROC diagram is depicted on the Fig. 2 for the four bands and first channel. Tab. IV shows results of the t-test of significance for each band, in the power spectrum of prediction error, originating from the first channel.

	p	t	df	s
δ-band	0.0944	-1.7274	30	7.4846
θ-band	0.0164	-2.5439	30	2.8343
α-band	0.1074	-1.6596	30	1.5903
β-band	0.0261	2.3399	30	2.7321

Tab. II *The significance of differences in the power spectrum of EEG signal (t-test) for 1st channel.*

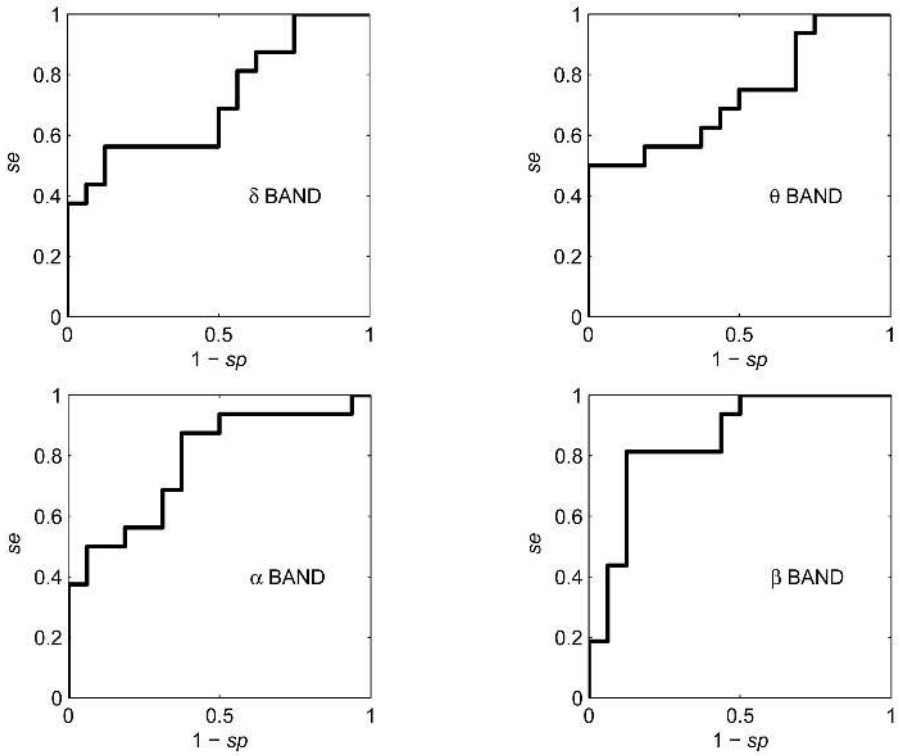

Fig. 2 *ROC for the 1st channel and $\delta, \theta, \alpha, \beta$ bands in the case of model error.*

ch	δ	θ	α	β
1	2.4×10^{-2}	6.4×10^{-3}	7.2×10^{-3}	1.7×10^{-2}
2	3.1×10^{-2}	2.3×10^{-2}	2.2×10^{-2}	1.7×10^{-3}
3	6.7×10^{-1}	4.9×10^{-1}	2.8×10^{-3}	1.7×10^{-2}
4	4.0×10^{-1}	7.7×10^{-2}	4.8×10^{-3}	2.1×10^{-2}
5	2.1×10^{-1}	5.5×10^{-2}	1.8×10^{-2}	5.0×10^{-3}
6	6.3×10^{-1}	1.9×10^{-1}	7.7×10^{-3}	6.6×10^{-3}
7	7.2×10^{-1}	4.4×10^{-1}	3.6×10^{-3}	2.0×10^{-2}
8	8.1×10^{-1}	9.0×10^{-1}	$\mathbf{5.1\times10^{-4}}$	1.9×10^{-1}
9	5.6×10^{-1}	3.0×10^{-1}	1.2×10^{-3}	2.0×10^{-2}
10	2.1×10^{-1}	5.2×10^{-2}	1.7×10^{-3}	3.6×10^{-3}
11	1.6×10^{-1}	7.6×10^{-2}	$\mathbf{7.1\times10^{-4}}$	1.6×10^{-2}
12	9.0×10^{-1}	4.0×10^{-1}	2.1×10^{-2}	5.8×10^{-2}
13	2.7×10^{-1}	3.7×10^{-1}	2.8×10^{-3}	2.0×10^{-3}
14	7.6×10^{-1}	6.1×10^{-1}	$\mathbf{4.1\times10^{-4}}$	1.0×10^{-3}
15	9.9×10^{-1}	3.9×10^{-1}	$\mathbf{9.4\times10^{-4}}$	$\mathbf{6.6\times10^{-4}}$
16	4.3×10^{-1}	5.6×10^{-1}	$\mathbf{4.5\times10^{-4}}$	1.4×10^{-3}
17	6.1×10^{-1}	3.9×10^{-1}	1.3×10^{-3}	1.2×10^{-3}
18	9.8×10^{-1}	9.1×10^{-1}	9.5×10^{-3}	$\mathbf{1.6\times10^{-4}}$
19	3.7×10^{-1}	4.4×10^{-1}	5.0×10^{-3}	$\mathbf{3.0\times10^{-4}}$

Tab. III *P-values for significant differences (AD × CN) in the case of model error (t-test).*

	p	t	df	s
δ-band	0.0238	-2.3815	30	1.5916
θ-band	0.0064	-2.9289	30	1.5619
α-band	0.0072	-2.8850	30	2.0161
β-band	0.0017	3.4542	30	3.9116

Tab. IV *The significance of differences in the power spectrum of the prediction error (t-test) for 1st channel.*

6. Conclusions

Band-power spectrum of raw EEG is an efficient tool for the classification of Alzheimer's diseased patients. The effectiveness of this approach can be measured by p-value of tested hypotheses. Bound-power spectral analysis of prediction error come to statistically significant results. Namely β-band relative power in the case channels 1, 2, 5, 6, 10, 13–19 has p-value < 0.01 in the case of two-sample two-sided t-test. The relative power of ANN prediction error is significantly lower in the case of Alzheimer's disease. It corresponds with hypothesis of decreased β-activity in the right frontal region of the human brain in the case of dementia. For classification purposes, β-band well as α-band, give more stable results in the case of channels 13–19 (in parietoocciptial region). This approach brings results, which are comparable with standard biomedical analysis [2], [3], [4].

Acknowledgement

This work created with support from CTU in Prague grant SGS11/165/OHK4/3T/14.

References

[1] Bishop C. M.: Neural Networks for Pattern Recognition, Oxford University Press, 1995.

[2] Revett K.: On the Use of Rough Sets for Artefact Extraction from EEG Datasets, Frontiers in the Convergence of Bioscience and Information Technologies, 2007, pp. 425-430.

[3] Castellaro C., et. al.: An Artificial Intelligence Approach to Classify and Analyses EEG Trace, Neurophysiol Clin, 2002, pp. 193-214.

[4] Tsoi A. C., et. al.: Classification of Electroencephalogram Using Artifical Neural Networks, University of Queensland St. Lucia, 1993.

[5] Petrosian A. A., et. al.: Recurrent Neural Network-Based Approach for Early Recognition of Alzheimer's Disease in EEG, Neurophysiol Clin, 2001, pp. 1378-87.

[6] Siegel J. H.: The Neural Control of Sleep and Waking, Springer, 2002.

[7] Lehmann C., et. al.: Application and Comparison of Classification Algorithms for Recognition of Alzheimer's Disease in Electrical Brain Ectivity (EEG), Journal of Neuroscience Methods, 2007, pp. 342-350.