NICOLAUS COPERNICUS UNIVERSITY IN TORUŃ
Faculty of Physics, Astronomy and Informatics

NOVA LINCS
LABORATORY FOR COMPUTER SCIENCE AND INFORMATICS

Analysis of brain activity based on long-term averaging of EEG signals

K. TOŁPA[2], Ł. FURMAN[1], E. RATAJCZAK[2], M. HAJNOWSKI[3], L. ALEXANDRE[4], W. DUCH[1]

1 Department of Informatics, Nicolaus Copernicus University, Toruń, Poland
2 Institute of Psychology, Faculty of Philosophy and Social Sciences, Nicolaus Copernicus University, Toruń, Poland
3 Department of Cognitive Science, Faculty of Philosophy and Social Sciences, Nicolaus Copernicus University, Toruń, Poland
4 NOVA LINCS and Universidade da Beira Interior, Covilhã, Portugal

Introduction

We investigate the diagnostic utility of resting-state EEG power maps based on long-term temporal averaging of signals in each channel, preserving spatial resolution. Two types of information derived from EEG measurements are important: power and dynamics. Our approach ignores dynamics, focusing on the asymptotic distribution of power that shows relative activations of different brain regions. To obtain the average power for each channel (electrode or source) we calculate Fourier spectra for second-long time windows, starting at each sample. The resulting power distribution maps (called **avPP**) are more detailed than those provided by the four microstate classes. Our approach allowed for the classification of schizophrenia patients based on data of rather poor quality (16 electrodes, low sampling rate) with accuracy comparable to sophisticated methods based on deep learning, and recurrence analysis. The power distribution maps are relatively stable for each individual and show relative stability for control vs. schizophrenia patients. The hope is that such simple method may be useful in clinical settings.

Data

- EEG recordings of 45 adolescent boys (10-14 years old) diagnosed with schizophrenia (schizotypical and schizoaffective disorders), described in Borisov et al. (2005).
- The control group consisted of 39 healthy schoolboys of similar age.
- Recorded in a wakeful relaxed state, with the eyes closed, using 16 electrodes placed according to the standard 10–20 system at O1, O2, P3, P4, Pz, T5, T6, C3, C4, Cz, T3, T4, F3, F4, F7, and F8.
- The sampling rate was 128 Hz.
- Only artifact-free EEG segments of the recordings were used for analysis.

Method

For each subject:

- given raw data matrix $U_k = (u_{ik})$, where the index i =1..N enumerates time-series samples and index k refers to electrodes (data streams, input channels),
- segment the data defining time windows with τ+1 samples, $w_k(t_i) = [u_{ik}, u_{ik+\tau}]$. These windows may have about one second.
- For each i=1..N-τ calculate STFT power spectrum, $S_k(t_i)$. Calculate the average total power P_k for channel k, sum all $S_k(t_i)$, and divide by the number of samples N-τ.
- If the sampling frequency is high windows may be shifted by several samples to speed calculations.
- Use the P vector of each subject for classification, and plot power maps.

Classification

Cumulative Average Power (avPP):

- The last cumulative average power values (after the summation over the whole time of the signal), and corresponding variances were used.
- Normalized within the frequency range for each subject separately and taken to create a feature matrix.
- We have tested several combinations of feature sets, selecting values for different bands (see Table).
- The classification was made using a linear SVM method implemented in the Scikit-learn Python library, with the stratified 5-fold cross-validation procedure.
- We used a recursive feature elimination with 10-fold cross-validation taken from the Scikit-learn function to exclude irrelevant features.

STFT Recurrence Analysis (STFT-RQA):

- Recurrence matrices (Marwan et al., 2007) are calculated directly from the STFT matrix, using Euclidean distance with a threshold (neighbourhood) parameter ε equal to the 35th percentile of the distance distribution.
- To quantify the dynamics within the system, the recurrence quantification analysis (RQA) features were used (Marwan et al., 2007).
- Features used for classification: entropy of diagonal lines (Lentr), trapping time (TT), average diagonal line length (L), laminarity (LAM), determinism (DET), longest vertical line length (Vmax), entropy of vertical line (Ventr), average white vertical line length (W), entropy of white vertical lines (Wentr).
- The chosen features were normalized.
- Irrelevant features were removed using recursive feature elimination in a 10-fold cross-validation procedure.

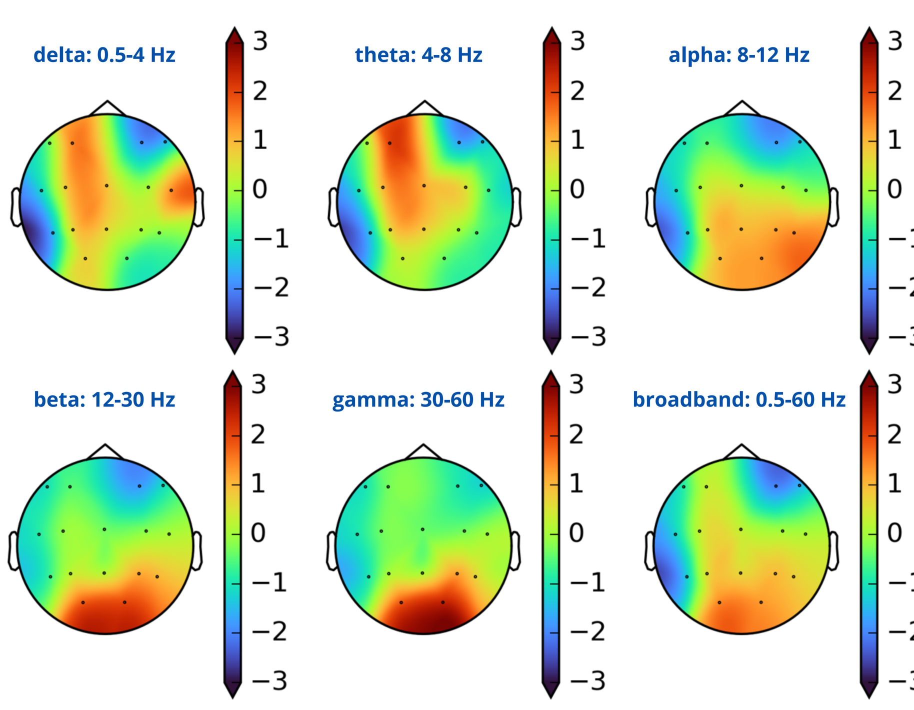
Microstates:

- Microstate analysis of EEG data focuses on the quickly changing spatial distribution of brain electric activity (scalp electrical potentials, Poulsen et al., 2018).
- Microstates are defined as transient quasi-stable spatial configurations usually lasting 60-120 ms (Michael & Koenig, 2018; Lehmann, 1990).
- A few prototypical topographies ('microstate classes') recur in time, representing quasi-simultaneous activity among the nodes of one of the large-scale networks; network activation is non-overlapping (Michael & Koenig, 2018).
- A fixed number of 4 microstate classes (Koenig et al., 1999) was calculated using the global approach and two clustering methods: TAAHC and modified k-means. The prototype topographies are presented in Fig. C.
- For each microstate class, six parameters were used for classification: occurrence, duration, coverage, global explained variance, mean spatial correlation, and global field potential.
- Relevant features were selected via recursive feature elimination using cross-validation.

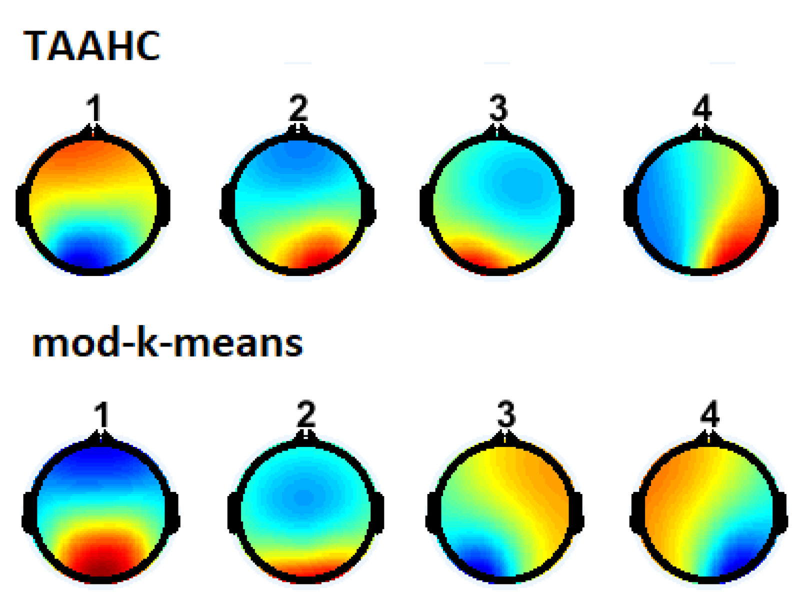
Convolutional Neural Network (CNN):

- A popular CNN approach considers the EEG data of each subject as an image.
- The images were created from NxN distance matrices from the 16 channel data, where the (i,j) pixel is given by d(x(i), x(j)), with i,j=1,...,N, where N is the number of EEG samples (128Hz x 60s), x(i) is a 16-d vector with each EEG channel at instant i, and d() is the Euclidean distance function.
- Image matrix was subsampled to 240x240 pixels. An example of such a matrix can be seen in Fig. D.
- A convolutional 3-layer CNN was constructed, with 16 3x3 filters in the first layer, 32 3x3 filters in the second and 64 3x3 filters in the third convolutional layer, ReLU layer after all convolutional layers, and the first of two fully connected layers (output layer is followed by a softmax), maxpooling layers after the ReLU layers of the convolutional blocks, and batch normalization layers after the second and third convolutional layers (with a total of 945170 parameters). It was trained using the Adadelta optimizer, for 200 epochs in each fold.

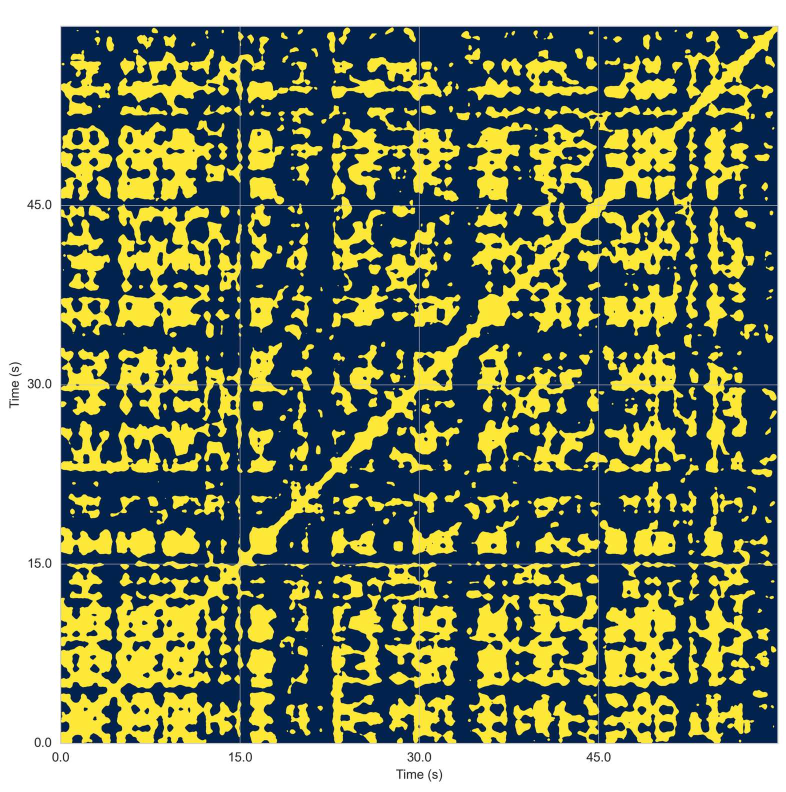
(A)



(C)



(B)



(D)

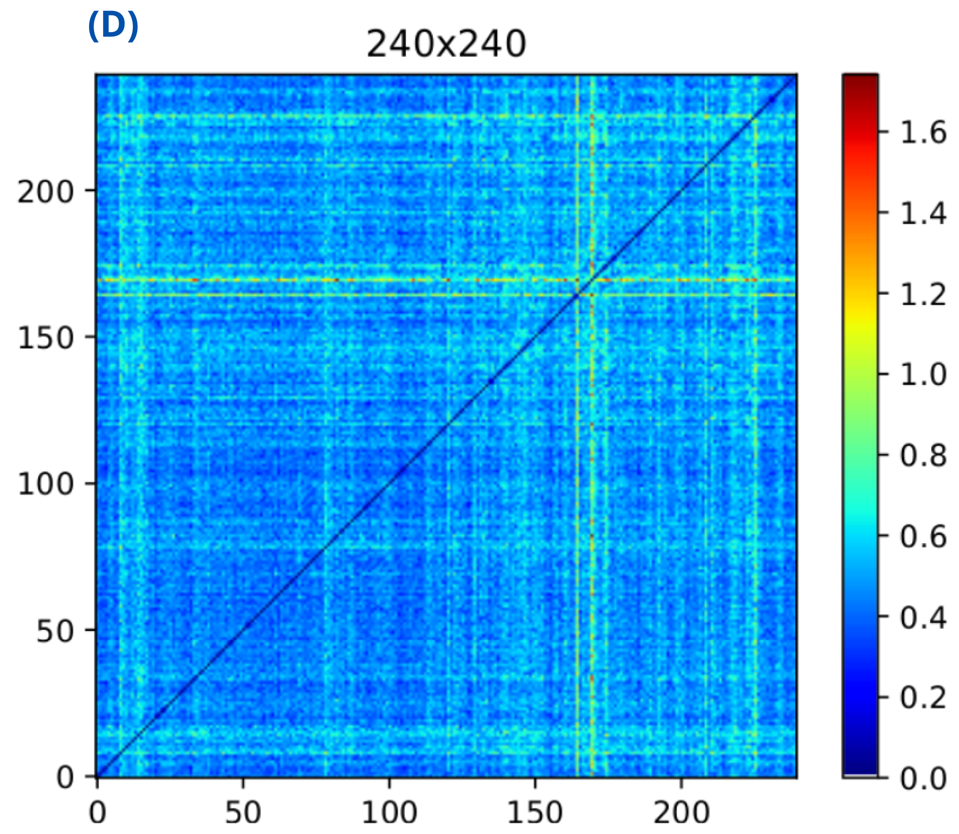


Figure 1 Examples of plots used in this study: (A) topographic maps of the cumulative average power values, (B) recurrence plot, (C) topographic maps of microstate prototypes, (D) distance matrix used by the CNN.

Results

Table 1

Summary of the best classification results for the features from RQA, microstate analysis and CNN

Method	RFECV	N dimensions	Mean Accuracy %			Variance of accuracy %		
			Total	Norm	Schizo	Total	Norm	Schizo
STFT RQA	yes	16	82.2	82.1	82.2	14.8	15.7	15.6
CNN	no	-	79.8	-	-	5.9	-	-
Microstates TAAHC	yes	3	78.5	69.3	86.7	17.2	23.3	12.2
Microstates K-means	yes	7	70.2	48.2	88.9	22.1	25.8	10.6

Table 2

Summary of the best classification results for the cumulative average power values. RFECV was used in all cases.

Window size	Bands	Variance of cumulative avg powers included	N dimensions	Mean Accuracy %			Variance of accuracy		
				Total	Norm	Schizo	Total	Norm	Schizo
256	delta-gamma	no	42	86.9	87.5	86.7	11.9	11.8	12.2
	delta-band	no	38	79.7	80	80	16.5	16.4	15.6
	delta-alpha	yes	48	83.3	87.5	80	13.7	11.1	12.8
	delta-theta	yes	27	84.7	80	88.9	12.9	15	9.4
	theta-beta	yes	6	77.4	70	84.4	17.5	18.6	13.3
	theta-alpha	no	26	77.4	80	75.6	18.4	16.4	19.4
	theta & beta	no	13	85.7	92.5	80	12.9	6.8	17.2
	broadband	yes	9	72.6	72.5	73.3	19.9	17.5	17.2

Acknowledgment and References

Supported by the National Science Center, grant UMO-2016/20/W/NZ4/00354.

Borisov, S. V., Kaplan, A. Ya., Gorbachevskaya, N. L., & Kozlova, I. A. (2005). Analysis of EEG Structural Synchrony in Adolescents with Schizophrenic Disorders. Human Physiology, 31(3), 255–261.
Furman, Ł., Duch, W., Minati, L., & Tołpa, K. (2022). Short-time Fourier transform and embedding method for recurrence quantification analysis of EEG time series. *The European Physical Journal Special Topics*, 1-15.
Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in neuroscience*, 267.
Koenig, T., Lehmann, D., Merlo, M. C. G., Kochi, K., Hell, D., & Koukkou, M. (1999). A deviant EEG brain microstate in acute, neuroleptic-naive schizophrenics at rest. European Archives of Psychiatry and Clinical Neuroscience 249(4), 205–211.
Lehmann, D. (1990). Brain Electric Microstates and Cognition: The Atoms of Thought. Machinery of the Mind, 209–224. https://doi.org/10.1007/978-1-4757-1083-0_10
Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports, 438*(5-6), 237-329.
Michel, C. M., & Koenig, T. (2018). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. NeuroImage, 180, 577–593.
Poulsen, A. T., Pedroni, A., Langer, N., & Hansen, L. K. (2018). Microstate EEGlab toolbox: An introductory guide. BioRxiv 289850, University of Zurich., 289850, 1–30.
Rawald, T., Sips, M., Marwan, N. (2017): PyRQA - Conducting Recurrence Quantification Analysis on Very Long Time Series Efficiently. - Computers and Geosciences, 104, pp. 101-108.