Time-frequency analysis combined with recurrence quantification for classification of onset of dementia using data from the oddball BCI paradigm

Krystian Dereziński
Academia Rerum Socialium,
Nicolaus Copernicus University,
87-100 Toruń, Poland
neuro@doktorant.umk.pl

Krzysztof Tołpa
Faculty of Philosophy and Social Sciences,
Nicolaus Copernicus University,
87-100 Toruń, Poland
ktolpa@umk.pl

Łukasz Furman
Dept. of Informatics, Faculty of Physics, Astronomy and Informatics,
Nicolaus Copernicus University,
87-100 Toruń, Poland
lfurman@doktorant.umk

Tomasz M. Rutkowski
Cognitive Behavioral Assistive Technology Team
RIKEN AIP
Tokyo, 103-0027 Japan
tomasz.rutkowski@riken.jp

Włodzisław Duch
Dept. of Informatics, Faculty of Physics, Astronomy and Informatics,
Nicolaus Copernicus University,
87-100 Toruń, Poland
wduch@umk.pl

Proc. of the 2022 Joint 12th International Conference on Soft Computing and Intelligent Systems, and 23rd International Symposium on Advanced Intelligent Systems (SCIS/ISIS 2022), Ise-Shima, Mie, Japan. IEEE Press, p. 1-6.

Abstract— **Reliable classification of EEG data based on a low number of electrodes is of great practical importance. Brain reactions to sensory stimuli may serve as digital biomarkers for detecting and monitoring the progression of dementia. Traditional approaches to the analysis of event related potentials (ERP) based on amplitude and latency variations do not have sufficient sensitivity to provide reliable biomarkers. Searching for a better approach we have combined time-frequency (TF) spectral representation of the EEG signal with recurrence quantification analysis (RQA). The non-linear features derived from the recurrence plots constructed in this way help to discriminate subtle differences of EEG signals. Auditory and visual stimuli were used in a single-trial oddball-type BCI experiments. Simple EEG using 16 electrodes was sufficient to achieve state-of-art classification accuracy of auditory and visual stimuli, using linear SVM method with selected RQA features. The differences between the standard linear amplitude ERP analysis and TF-RQA approach is also shown in the histograms of the SVM linear projection and in the Uniform Manifold Approximation and Projection (UMAP) plots. The small scale of these experiments and lack of longitudinal observations does not allow justifying the hypothesis linking mild cognitive impairment (MCI) with auditory response degradation, but the TF-RQA should be a good tool in creation of such biomarkers.**

Keywords—Recurrence Quantification Analysis, Event Related Potential, Short-Time Fourier Transform, Takens embedding theorem, EEG Time Series

I. INTRODUCTION

In the highly developed countries dementia affects about 5% of the population. Alzheimer's Disease International (ADI) reports that there were about 55 million people living with dementia in 2022 around the world, and there are almost 10 million new cases each year. With the aging population, the number of dementia cases is expected to almost double in the next 10 years, reaching 78 million in 2030 [1]. Initial symptoms of mild cognitive impairments (MCI), related to memory loss, problems with the use of language, visuospatial perception, or executive function, may signal the slow onset of Alzheimer's disease, the most common form of dementia that accounts for 60-70% of all such cases [1]. Dementia is currently the seventh leading cause of death among all diseases, and one of the major causes of disability and dependency among older people globally. It has significant social and economic consequences in terms of direct medical and social care costs and the cost of informal care [1]. In 2019, the estimated total social cost of dementia worldwide was 1.3 trillion USD, and it is expected to exceed 2.8 trillion USD by 2030 [1].

The causes of dementia still remain unclear, despite many theories that involve genetic and cellular processes, influenced by a lifestyle and environmental factors [2]. Early diagnosis is important because cognition-based training in people with MCI may improve their memory, executive function, attention, and general mental performance [3]. Therefore, development of AI-based applications to support early screening for MCI, combined with various interventions designed to maintain healthy cognition, is of utmost importance.

Brain activity, expressed in the oscillations of electric potentials generated by neural populations, changes in response to the sensory stimulation and the internal neurodynamics [4]. Electroencephalography (EEG) is a cost-effective, non-invasive technique that measures electrical activity of the brain, providing in vivo data with high temporal resolution. EEG has been used successfully in the study of neurocognitive disorders, such as epilepsy [4], schizophrenia [5], autism [6], Alzheimer disease [7] and many other disorders.

Interpretation of noisy, non-stationary EEG signals is difficult [8]. Linear amplitude analysis of event-related potentials (ERPs) are the most reliable technique, based on averaging EEG signals arising in response to repetitive sensory or visual stimuli. Averaging signals over many trials removes background noise, increasing signal-to-noise-ratio, exposing relatively stable electrophysiological responses to identical sensory stimuli [9]. Plotting changes of electric potential as a function of time after the onset of stimuli shows several components, such as decrease after 100 ms (N100), increase starting after 300 ms (P300), and later N400 and P600 components seen in language processing experiments [10]. Averaging a large number of trials that differ by latency and amplitude does not allow for observation of true variations of the neural response, effects of attention, priming, habituation, or fatigue [8]. Methods for single-trial ERP analysis based on regularization of linear discriminant analysis (LDA) by shrinkage filters are popular in applications to the brain-computer interfaces, but not in medical diagnostics, where frequently ERPs from a single electrode are analyzed.

We are interested in diagnostic methods that do not require large number of trials, and could provide detailed analysis of properties related to neurodynamics. Deep-learning methods applied to analysis of EEG gain popularity, but are still difficult to interpret [11]. One of the promising, and still little explored approach, is based on analysis of recurrence patterns, that provides non-linear interpretable features [6]. Analysis of time-frequency (TF) spectrograms, illustrating changes in time of distribution of power in different frequencies, is another good approach to learn EEG structure [12]. In this paper, we combine both approaches, using TF signal representation (power spectra for each time point) with similarity evaluation of such vectors to create recurrence plots and extract non-linear features for classification using RQA, hence the TF-RQA acronym.

Recurrence plots (RPs) are visual representation of recurrence matrices, that store information about similarity of the current state $\mathbf{X}(t)$ of the system to the previous states. One can define recurrence using some distance metrics $\mathbf{D}(t,t')=||\mathbf{X}(t)-\mathbf{X}(t')||$, or use tolerance threshold ε (usually related to the noise in the system) and a step function to create binary matrices $\mathbf{R}(t,t')=(\mathbf{\Theta}(\varepsilon-||\mathbf{X}(t)-\mathbf{X}(t')||)$. If the current trajectory is closer than ε to the previous one element of the recurrence matrix is equal to 1, otherwise it is 0. Other possibilities may use instead of trajectory points $\mathbf{X}(t')$ a few optimized reference points $\mathbf{X}(t_k)$, allowing for reduction of dimensionality of the state vectors, $\mathbf{Y}(t,t_k)=||\mathbf{X}(t)-\mathbf{X}(t_k)||$. This approach creates fuzzy partition of the state space, leading to Fuzzy Symbolic Dynamics [3, 13].

Recurrence Quantification Analysis (RQA) defines a number of features that can be derived from the binary recurrence matrices [9]. Recurrence rate RR is the percentage of 1 bits in the recurrence matrix (or black dots in the recurrence plots RPs). Vertical lines are used to create 3 features: trapping time TT is calculated as the average length of the vertical lines, laminarity LAM is the percentage of recurrence points which form vertical lines, and Vmax is the length of the longest vertical line. Diagonal lines also contribute 3 features: determinism DET is the percentage of recurrence points which constitute diagonal lines, AvgL is the average diagonal line length, and Lmax the longest diagonal line length. Other features include various entropy measures (classical Shannon entropy, information entropy) that should be calculated in an unbiased way and applied to the RQA features [14]. In contrast to other non-linear analysis methods, RQA can be applied to relatively short non-stationary time series.

Andrade et al. [8] have used recurrence analysis in a small study (11 subjects) to distinguish rare and frequent tones in an oddball study. They have calculated ERPs and classified responses to tones using linear amplitude (LA) analysis of P300 component. Takens embedding theorem requires selection of embedding dimension m and delays τ to sample original signal at times $[x(t), x(t+\tau), \dots x(t+(m-1)\tau)]$. Since optimal values of these parameters change between trails and subjects, they have averaged them for all values from 1 to 20. Using different embeddings of the original EEG signals and fixing the number of neighboring states 6 recurrence features were calculated and in this space classification performed. Results of RQA discrimination of responses to different tone types in this oddball experiment were based on AvgL values for the Pz electrode. They have not been significantly different from the linear amplitude analysis. However, RQA features were calculated from the single trials, and values averaged later for rare and frequent tones.

Our approach differs in many aspects. In this study, we probed the potential of the Recurrence Quantification Analysis to improve the characterization of single trial EEG responses. Below we present results of the RQA analysis of ERPs from BCI experiments using 16-channel EEG system and signal representation based on Short-Time Fourier Transformation (STFT) power spectra, instead of the standard signal embedding approach used in EEG signal classification. This simplifies interpretation, as each state is simply a power spectrum created from trials of 1200 ms lengths, and recurrence matrices show similarity of these spectra. The goal is to recognize ERP signals resulting from the auditory vs. visual stimuli.

II. MATERIALS AND METHODS

A. EEG acquisition and preprocessing

The EEG dataset used in this paper has been provided by one of the current project's co-authors and a member of a previous study [15] conducted with 16 BCI-naïve subjects (mean age 21.8 with a standard deviation of 0.75). All the original experiments, first reported in [15], were performed at the Life Science Center of TARA, University of Tsukuba, Japan. The online EEG BCI experiments were performed following the WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. The subjects of the experiments received financial gratification.

The 200 ms long unimodal (visual or auditory) stimuli were delivered from five different spatial locations. In the case of visual speller paradigms, large-size hiragana characters were flashed one at a time on a big computer display positioned in front of the subject (as it is usually offered in the oddball-based P300 visual speller). In the case of the auditory modality, a sound source was positioned at a spatial location congruent with a visual letter (set at -45°, -22.5°, 0°, 22.5°, and 45° in front of the head). The subjects were instructed to spell five random sequences of hiragana letters presented visually or audibly in each session. Each target was presented ten times. Each subject first conducted a short psychophysical test with a button press response to confirm understanding of each modality's experimental set-up.

During the original online BCI experiments, as described in [15], the EEG signals were captured with 16 active electrodes using amplifier system g.USBamp by g.tec, with a sampling frequency of 512 Hz. The electrodes were placed at the following head locations Cz, CPz, POz, Pz, P1, P2, C3, C4, O1, O2, T7, T8, P3, P4, F3, and F4, as in the 10/10 extended international system. The ground and reference electrodes were attached at FCz and the earlobe. After EEG signal acquisition, time-series epochs were extracted within an interval of -200 to 1000 ms around each stimulus onset. Afterward, each epoch's and condition's data (visual and auditory) were preprocessed. Bad epochs were identified using the auto-reject extension for MNE software (Jas et al., 2016, 2017) with a global rejection criterion and dropped from the data. In the current project, we model dementia-related and delayed P300 responses [16, 17, 18, 19] using auditory BCI paradigm responses from the study mentioned above [18] and normal EEG using visual ERPs (non-delayed and with higher amplitude). A working hypothesis of the current study is to evaluate the possibility of successfully discriminating between two types of ERPs carrying standard (visual P300) versus delayed and with lower amplitude (auditory P300 modeling dementia).

B. *Methods*

For every electrode and each epoch acquired separately for every subject STFT vectors were computed in sliding time windows separately. For this purpose, we have used the TensorFlow implementation, defining Hamming window type with 240 samples (corresponding to 1488 ms), shifted by a single sample. This number of samples proved to be a good compromise, as shorter time frames will introduce uncertainty in the frequency determination of the Fourier transform, and larger windows increase uncertainty of time in which calculated spectra arise.

STFT analysis returns samples representing each vector for each time window, with the number of frequency bins nfft set to 512. This operation resulted in a matrix S containing a representation of time series based on STFT. These vectors have a clear interpretation, showing peaks of characteristic frequency.

Next, distance matrices were calculated from the obtained STFT vectors using Euclidean metric. For the recurrence plots the matrices were binarized applying similarity tolerance threshold ε equalled to the 35th percentile of the distance distribution.

Non-linear features were calculated from the recurrence matrices using the recurrence quantitative analysis (RQA). We have used our own implementation of recurrence quantification analysis, based on modification of the *recurrence_python* software [20]. The RQA features included quantities: recurrence rate, determinism, average diagonal line length, entropy diagonal lines, laminarity, trapping time, longest vertical line length, entropy vertical lines, average white vertical line length, entropy white vertival lines. At this point seperate RQA for each epoch, electrode, condition and subject were obtained. Next, values of the features were averaged between epochs for each electrode and condition in a similar manner as the event related potentials are obtained as the average potential calculated between epochs.

Since the recurrence rate feature was constant, for the selection of the threshold parameter value it was considered irrelevant and dropped for the further analysis.

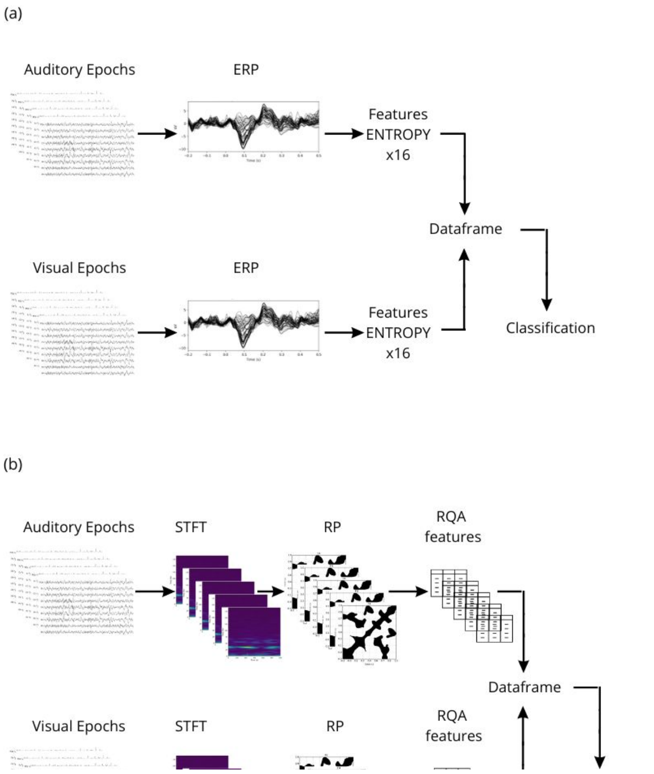

Fig. 1. LA(a) and RQA(b) pipeline. ERP - Event Related Potential, STFT - Short-Time Fourier Tranformation, RP - Rcurrence Plots, RQA - Recurrence Quantification Analysis.

The data from the remaining features was normalized and cleaned from outliers using Robust Chauvenet Outlier Rejection [21]. Furthermore, to exclude redundant electrodes and features, we have performed recursive feature elimination with 10-fold cross-validation using scikit-learn function. In every step of this process, the algorithm decreased the number of component from the data by one, based on the lowest absolute value of the SVM coefficient and performing cross-validation in the space of reduced dimensionality. The analysis resulted in 28 chosen components. Those feature components were then used as an input to the classifiers.

We have tested several classifiers: support vector machine (SVM) with linear kernel, and with quadratic and cubic kernels, Gaussian density Bayesian classifier that provides linear decision boundaries (LDA), and a Random Forest Classifier (RFC), all implemented in the popular scikit-learn package [22]. The predicted outcome was evaluated using leave-one test.

III. Results

An artificial intelligence and neurotechnology approach using machine learning to identify task load (by different modality), which was proposed as a model for cognitive responses for dementia [18], has resulted in an encouraging accuracy of rate of classification.

We obtained results from a dataset of sixteen subjects. The true statistic mean accuracy for 95% confidence interval was estimated using bootstrap procedure. The results (median of the mean distributions and upper and lower bounds) are summarized in Figure 2. and Table 1.

The best results were obtained for linear SVM - with accuracy rate of 100%. SVM with quadratic and cubic kernels give slightly worse results, with some variance, while Gaussian-based LDA and Random Forest Classifiers (RFC) are significantly worse.

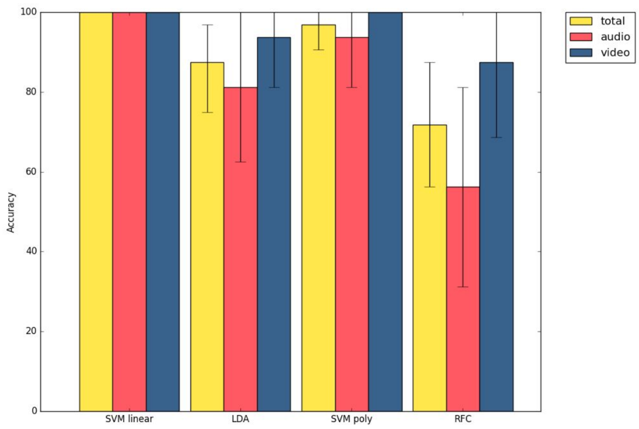

Fig. 2. Median cross-validation accuracies for all subject dataset of modality (auditory - visual) classification.

In Fig. 3 RQA vectors used for classification using linear SVM are projected using the SVM weight vector.

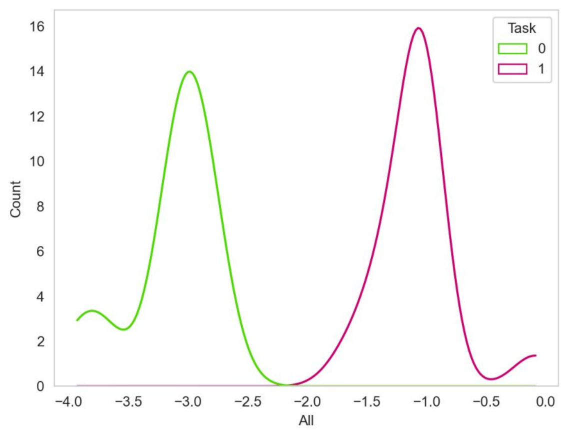

Fig.3. Histograms of the projection of 28 RQA feature values used for classification, selected by the random forest, for all subjects, and for all data.

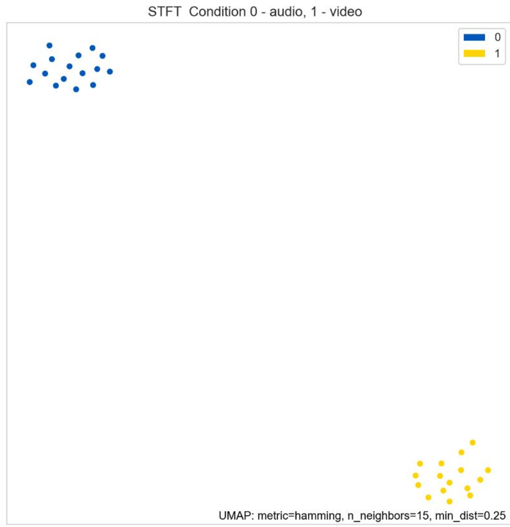

Fig. 4. UMAP visualization of the 28 dimensional vectors based on the STFT-RQA approach for all subjects.

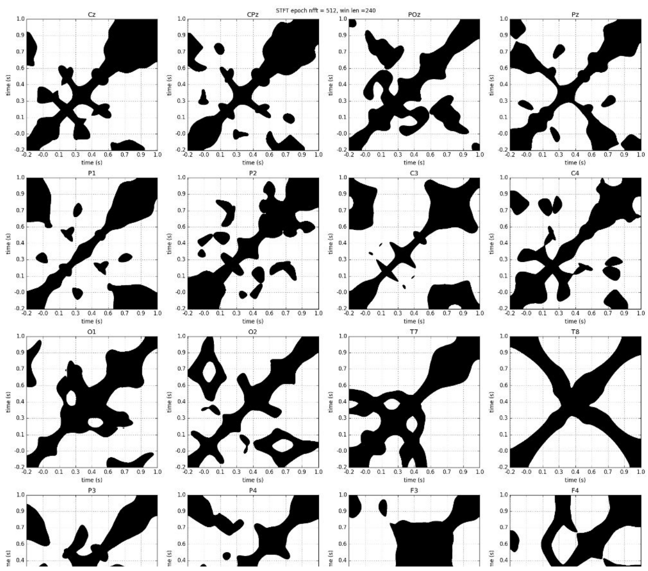

Fig. 5. Recurrence plots for **visual** condition and **target stimulus type** for electrodes Cz, CPz, POz, Pz, P1, P2, C3, C4, O1, O2, T7, T8, P3, P4, F3, and F4, as in the 10/10 extended international system.

Such visualization helps to distinguish cases that are quite distinct from those that are similar. In this case our method found representation that distinguishes all cases clearly. This is also confirmed by UMAP visualization (Fig. 4) that shows wide separation of clusters for auditory and visual stimuli.

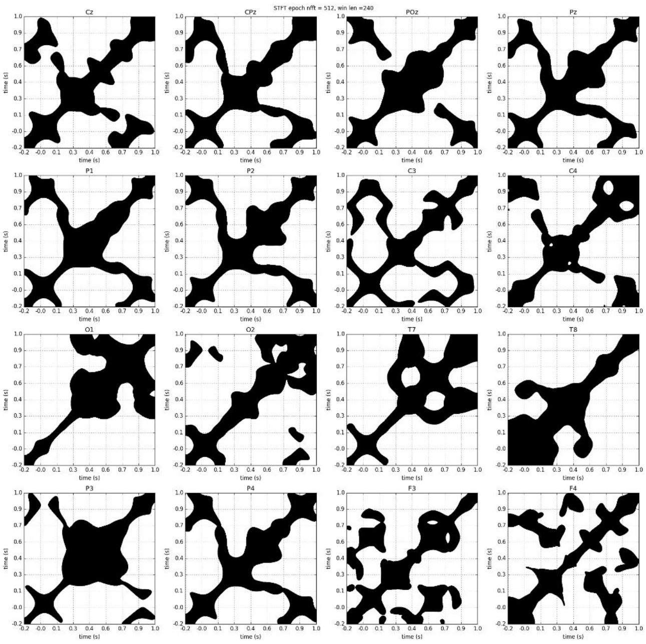

Fig. 6. Recurrence plots for auditory condition and target stimulus type for electrodes Cz, CPz, POz, Pz, P1, P2, C3, C4, O1, O2, T7, T8, P3, P4, F3, and F4, as in the 10/10 extended international system.

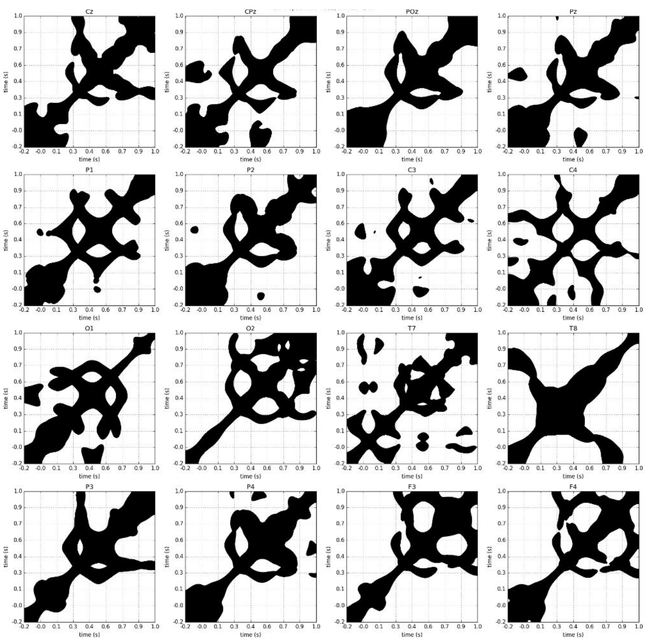

Fig. 7. Recurrence plots for **visual** condition and **nontarget stimulus type** for electrodes Cz, CPz, POz, Pz, P1, P2, C3, C4, O1, O2, T7, T8, P3, P4, F3, and F4, as in the 10/10 extended international system.

CONCLUSION

In this study we have tested the effectiveness of ERP classification using TF-RQA approach, a combination of short-time Fourier transformation (as in spectrograms) with recurrence quantification analysis. Results were compared with the standard method based on linear amplitude analysis (LA). Instead of focusing on ERP features, such as the P300, we have used non-linear features extracted from the recurrence plots (such as shown in Fig 5-7) that show similarity of power spectra at different points in time. These plots show ERP patterns measured by various electrodes, showing significant differences between visual and auditory stimuli that are expressed in RQA features.

Classification results for four methods are presented in Table 1. Linear SVM achieves in this case perfect 100% accuracy. For SVM with quadratic or cubic kernel we have 96.9%, LDA gave 87.5% and RFC 71.9%. where we find that for a low number of features like 28 RFC is not a optimal method.

Classification method	Mean Performance	
	Median	*95% CI*
SVM linear	100.0	[100, 100]
SVM poly	96.9	[90.6, 100]
LDA	87.5	[75.0, 96.9]
RFC	71.9	[56.2, 87.5]

Tab. 1. Accuracies for all subject dataset of modality (auditory - visual) classification. SVM - Support Vector Machines, LDA - Linear Discriminant Analysis, RFC - Random Forest Classification

RQA combined with linear SVM provided a perfect distinction between the responses to the two main modalities in the oddball experiment (visual and auditory). Linear amplitude analysis of ERPs has also high discrimination power, reaching almost 90% in the leave-one-out tests, with auditory stimuli discrimination being slightly more difficult. Our initial hypothesis was that such experiments may help to distinguish people at the onset of MCI problems, when the auditory perception starts to fail, and visual perception still works fine. However, to test this hypothesis, more complex tasks with larger number of participants are necessary. In this pilot experiment, we can only point to the accuracy of STFT signal representation combined with recurrence analysis in analysis of such data.

RQA can extract a different type of information from the EEG signal, as suggested by the theoretical basis of RQA. Despite higher computational demands and conceptual complexity of the method, RQA should be preferred over amplitude analysis (LA) to discover features that can help in more detailed characterization of brain responses to event-related stimulation. RQA can be used as a complementary tool in the analysis of single samples of electrophysiological data. In particular, analysis of pooled EEG data with other modalities that provide additional variables (such as psychometric tests or fMRI) may be useful for elucidating various neural correlates of measures of amplitude and complexity, such as provided by RQA.

Acknowledgments. Support by the Polish National Science Center grant, UMO-2016/20/W/NZ4/00354, is gratefully acknowledged.

REFERENCES

[1] Lynch, C. (2020). World Alzheimer Report 2019: Attitudes to dementia, a global survey: Public health: Engaging people in ADRD research. Alzheimer's & Dementia, 16, e038255.

[2] Sprague, B. N., Freed, S. A., Webb, C. E., Phillips, C. B., Hyun, J., & Ross, L. A. (2019). The impact of behavioral interventions on cognitive function in healthy older adults: A systematic review. Ageing Research Reviews, 52, 32–52.

[3] Duch, W., & Dobosz, K. (2011). Visualization for understanding of neurodynamical systems. Cognitive Neurodynamics, 5(2), 145–160.

[4] Sharmila, A. (2018). Epilepsy detection from EEG signals: A review. Journal of Medical Engineering & Technology, 42(5), 368–380.

[5] Wang, F., Hujjaree, K., & Wang, X. (2021). Electroencephalographic Microstates in Schizophrenia and Bipolar Disorder. Frontiers in Psychiatry, 12.

[6] Bosl, W. J., Tager-Flusberg, H., & Nelson, C. A. (2018). EEG Analytics for Early Detection of Autism Spectrum Disorder: A data-driven approach. Scientific Reports, 8(1), 6828.

[7] Ouchani, M., Gharibzadeh, S., Jamshidi, M., & Amini, M. (2021). A Review of Methods of Diagnosis and Complexity Analysis of Alzheimer's Disease Using EEG Signals. BioMed Research International, 2021, e5425569.

[8] Andrade, K. C., Wehrle, R., Spoormaker, V. I., Sämann, P. G., & Czisch, M. (2012). Statistical evaluation of recurrence quantification analysis applied on single trial evoked potential studies. Clinical Neurophysiology, 123(8), 1523-1535.

[9] Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. Physics Reports, 438(5), 237–329.

[10] Horvath, A., Szucs, A., Csukly, G., Sakovics, A., Stefanics, G., & Kamondi, A. (2018). EEG and ERP biomarkers of Alzheimer's disease: a critical review. Frontiers in bioscience (Landmark edition), 23, 183-220.

[11] Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. Journal of neural engineering, 16(3), 031001.

[12] Tawhid, M. N. A., Siuly, S., Wang, H., Whittaker, F., Wang, K., & Zhang, Y. (2021). A spectrogram image based intelligent technique for automatic detection of autism spectrum disorder from EEG. PLOS ONE, 16(6), e0253094.

[13] Dobosz, K., & Duch, W. (2010). Understanding neurodynamical systems via fuzzy symbolic dynamics. Neural Networks, 23(4), 487–496.

[14] Leonardi, G. (2018). A Method for the computation of entropy in the Recurrence Quantification Analysis of categorical time series. Physica A: Statistical Mechanics and Its Applications, 512, 824–836.

[15] M. Chang, N. Nishikawa, Z. R. Struzik, K. Mori, S. Makino, D. Mandic, and T. M. Rutkowski, "Comparison of P300 responses in auditory, visual and audiovisual spatial speller BCI paradigms," in Proceedings of the Fifth International Brain-Computer Interface Meeting 2013, (Asilomar Conference Center, Pacific Grove, CA USA), pp. Article ID: 156 (1–2), Graz University of Technology Publishing House, Austria, June 3-7, 2013.

[16] Jas, M., Engemann, D., Raimondo, F., Bekhti, Y., & Gramfort, A. (2016, June). Automated rejection and repair of bad trials in MEG/EEG. In *2016 International Workshop on Pattern Recognition in Neuroimaging (PRNI)* (pp. 1-4). IEEE.

[17] Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage*, *159*, 417-429.

[18] Rutkowski, T. M., Koculak, M., Abe, M. S., & Otake-Matsuura, M. (2019, May). Brain correlates of task–load and dementia elucidation with tensor machine learning using oddball BCI paradigm. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8578-8582). IEEE.

[19] Rutkowski, T. M., Zhao, Q., Abe, M. S., & Otake, M. (2018). AI Neurotechnology for Aging Societies--Task-load and Dementia EEG Digital Biomarker Development Using Information Geometry Machine Learning Methods. arXiv preprint arXiv:1811.12642.

[20] Reis, B. Scientific software written in Python 3 for the creation of Recurrence Plot (RP), Cross Recurrence Plot (CRP), and Joint Recurrence Plot (JRP). It also allows the quantification analysis of RP. https://github.com/bmfreis/recurrence_python

[21] Maples, M. P., Reichart, D. E., Konz, N. C., Berger, T. A., Trotter, A. S., Martin, J. R., ... & Salemi, C. P. (2018). Robust chauvenet outlier rejection. The Astrophysical Journal Supplement Series, 238(1), 2.

[22] Hao, J., & Ho, T. K. (2019). Machine learning made easy: a review of scikit-learn package in python programming language. Journal of Educational and Behavioral Statistics, 44(3), 348-361.

.