2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Older Adult Mild Cognitive Impairment Prediction from Multiscale Entropy EEG Patterns in Reminiscent Interior Image Working Memory Paradigm

Tomasz M. Rutkowski[1,2,3], Masato S. Abe[1], Tomasz Komendziński[3], and Mihoko Otake-Matsuura[1]

***Abstract*— We discuss the practical employment of a machine learning (ML) technique within AI for a social good application. We present an application for elderly adult dementia onset prognostication. First, the paper explains our encouraging preliminary study results of EEG responses analysis using a signal complexity measure of multiscale entropy (MSE) in reminiscent interior working memory evaluation tasks. Then, we compare shallow and deep learning machine learning models for a digital biomarker of dementia onset detection. The evaluated machine-learning models succeed in the most reliable median accuracies above 80% using random forest and fully connected neural network classifiers in automatic discrimination of normal cognition versus a mild cognitive impairment (MCI) task. The classifier input features consist of MSE patterns only derived from four dry EEG electrodes. Fifteen elderly subjects voluntarily participate in the reported study focusing on EEG-based objective dementia biomarker advancement. The results showcase the essential social advantages of artificial intelligence (AI) application for the dementia prognosis and advance ML for the subsequent use for simple objective EEG-based examination.**

***Clinical relevance*— This manuscript introduces an objective biomarker from EEG recorded by a wearable for a plausible replacement of a mild cognitive impairment (MCI) evaluation using usual biased paper and pencil examinations.**

I. INTRODUCTION

A rise in dementia cases globally results in significant cost inflation in healthcare. Nearly 50 million older adults suffer from dementia-related neurocognitive maladies, as detailed by the World Health Organization (WHO) [1], and this number expects to triple through the subsequent three decades [2]. This increasing difficulty calls for a feasible utilization of AI to advance early diagnostics for subsequent cognitive well-being monitoring and preservation with so-called "digital pharma" or "beyond a pill" non-pharmacological-therapeutical (NPT) strategies [3]. An ultimate dementia determination is only possible by postmortem autopsy. A differential examination with other types of age-related brain neurodegeneration is usually ventured. A cognitive standing examination, such as the Montreal Cognitive Assessment (MoCA) [4], [5], is ordinarily utilized to quantify the severity of dementia. Objective medical imaging methods, such as functional magnetic resonance imaging (fMRI) [6] or EEG [7], [8], [9] together with behavioral measures [10], [11], are recently in ongoing expansion to provide an early onset of a mild cognitive impairment (MCI) prediction and subsequent monitoring. In this study, we test a hypothesis that reminiscent (childhood age) versus modern/contemporary interiors within a working memory paradigm, inside Western and Japanese designs, are helpful for a new EEG-based dementia biomarker application. We explicitly decide to test a wearable EEG in the current project to develop a home-based biomarker shortly. We select to use a popular wearable MUSE EEG system (InterAxon Inc., Toronto, Canada) that allows for a quick collection of EEG and behavioral data. The MUSE headband provides an acceptable and quantifiable event-related-potential (ERP) and broadband EEG collection as shown in [12], [13]. Dry-electrode-based EEG systems result in more noisy EEG signals comparing to clinical-grade systems. To deal with noisy time-series, we propose to utilize a signal complexity in the form of multiscale entropy (MSE), which is more robust to noise comparing to traditional time-domain features [14], [15], [12].

[1]Tomasz M. Rutkowski, Masato S. Abe, and Mihoko Otake-Matsuura are with Cognitive Behavioral Assistive Technology Team, RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan, tomasz.rutkowski@riken.jp

[3]Tomasz M. Rutkowski is with Graduate School of Education, The University of Tokyo, Tokyo, Japan.

[3]Tomasz M. Rutkowski and Tomasz Komendziński are with Department of Cognitive Science, Nicolaus Copernicus University, Toruń, Poland.

II. METHODS

We carry brainwave data recording experiments with older adults in the RIKEN Center for Advanced Intelligence Project (AIP). The study adheres to human subject experimental involvement guidelines and ethical review from the RIKEN Ethical Committee for Experiments with Human Subjects and The Declaration of Helsinki. In the study, 15 seniors (11 females; mean age of 74.3 years old; ± 6.4 years' standard deviation of age; recruited from Silver Human Resources Center and Honobono Laboratory, Japan, 10 subjects with MCI evaluation based on MoCA≤ 25) took part. All participants received monetary gratification for their participation in the study, and they gave informed written consent. All participants accepted financial gratification for their co-operation in the study, and they provided informed written consent. We use for EEG data collection a four-channel portable MUSE 2016 headband by InteraXon Inc., Canada. It has been shown already that the MUSE device allows for a reliable EEG capture from preset $AF7$,$AF8$,$TP9$, and $TP10$ electrode locations using dry electrodes and careful setting conditions [12]. The ground reference electrodes are set at the forehead. We prepare our in-house EEG capture environment in Python using muse-lsl [13] library, which

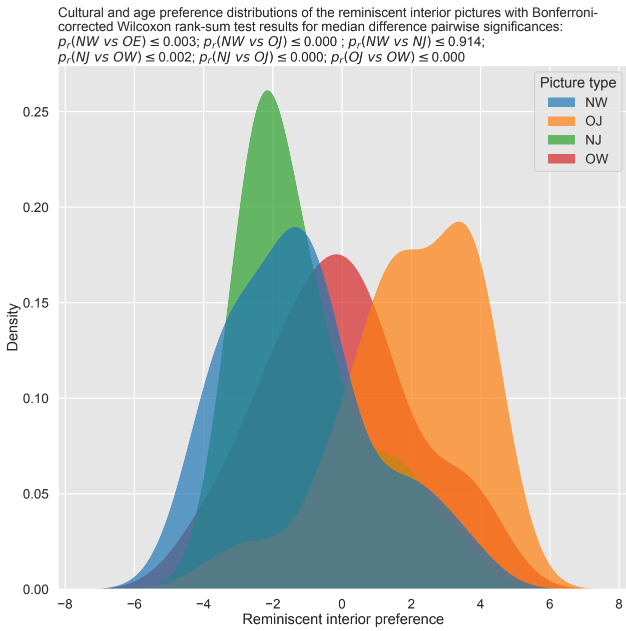

Fig. 1. Subject preference scores for the four groups in reminiscent (childhood age) versus modern/recent interiors within Western and Japanese designs. NW denotes the picture types for new Western, OW for old/reminiscent Western, NJ for contemporary Japanese, and OJ for old/reminiscent Japanese. The response distributions from all 15 participating subjects clearly show significantly lower preferences with $p_r(N*\ \ vs\ \ O*) \leq 0.05/4 = 0.0125$ as tested with Bonferroni-corrected Wilcoxon rank-sums tests, where $[*]$ stands for J or W, for the modern (NJ or NW) designs (blue and green distributions versus red and orange). At the same time, the reminiscent Japanese interiors (orange) were the more favorable. The reminiscent Western interiors (red) resulted in intermediate preference results with also statistically significant differences to the other groups.

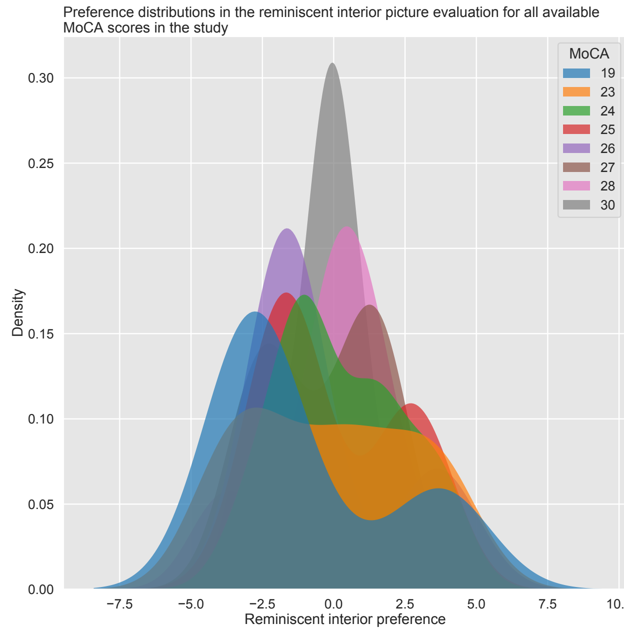

Fig. 2. Subject preference distributions depicted for all MoCA scores available in the project. Higher MoCA scoring subjects resulted in unimodal while the lower scoring participants in bimodal distributions.

communicates with MAX [16] visual programming environment for stimulus presentation and reminiscent preference responses recording. We disinfect each time the MUSE headband with alcohol wipes and place it on a subject's head with a careful checks for EEG average amplitudes not to exceed thresholds indicating EMG contamination. We use 50Hz notch filter to remove power line interference, and we further bandpass the signal in a range of $1 \sim 30$Hz to minimize out of band noise. Each elderly participant, after the experimental procedure and purpose explanations, signs an informed consent. There are eight types of interior pictures in each EEG recording session in Japanese and Western styles within old (post World War II) and modern (recent designs), creating categories with two examples from each. The pictures we present in an oddball-style paradigm. Before each random order presentation of eight interiors, we request the participant to memorize and focus on one of them and ignore the remaining seven. Altogether, each subject receives 64 visual stimulations. After the EEG session, we conduct a quick preference evaluation behavioral experiment using a touchpad slider. We ask the participants to evaluate a subjective reminiscence of each image on a linear scale from -5 to $+5$. We discuss the results of both experimental modalities in subsequent sections. We segment EEG into three-second long intervals, starting each stimulus onset time, in which the participants observed presented reminiscent interior images, with MNE version 0.23.0 package in Python [17]. Multiscale Entropy (MSE) is a tool allowing for quantification of a signal complexity at varying time scales [14]. MSE involves successive computations of a sample entropy [18] estimated on coarse-grained sequences representing system dynamics on different time scales [14]. Principally the MSE calculation consists of a coarse-graining or downsampling of the analyzed time-series – essentially looking at increasingly coarser time resolutions. In our project, we analyzed four-electrode EEG separately over the signal segmented into three-second long post-stimulus intervals using *neurokit2* library [19]. The resulting MSE features are summed sample entropy values over all analyzed scales within the analysis window of three seconds [15]. We test classifiers available in *the scikit-learn* library version 0.24.1 [20] for binary classification of MCI versus normal cognition of the 15 participants in our reminiscent interior image working memory study using input MSE values from four EEG channels only as feature vectors. We use a ten-fold cross-validation procedure due to a limited number of available subjects, with a chance level of 66% due to a more significant number of MCI versus normal subjects available in our study. We test the following classification methods with input feature standard scaling using a removal a mean and division by a variance as follows: **a logistic regression (LR)** w; with subsequent *liblinear* solver; a maximum iteration count set to 1000; **a linear discriminant analysis (LDA)** with a least-squares solver and no shrinkage; **a support vector machine with the linear kernel (linearSVM)**, with a loss function set to a squared hinge and $l2-$penalty linear kernel; **support vector machine with a radial basis function (rbfSVM)** using a kernel coefficient *gamma* equal to $1/4$ (rendering an inverse of feature-length); **support vector machine with a polynomial kernel (polySVM):** with a second-degree polynomial kernel employing a coefficient *gamma* set to further to $1/4$ and an independent term in kernel function $coef0 = 1.0$. **support vector machine with a sigmoid kernel (sigmoidSVM)** with a kernel coefficient *gamma* set also to $1/4$; **a random forest classifier (RFC)** adopting a number of trees in the forest equal to 200, a split criterion by mean squared error used, without a maximum tree depth limitation, and a number 2 set as a minimum number of samples required for a split; **a fully connected deep neural network (FNN)** applying densely connected layers with rectified linear units (ReLU), with architecture using a four-unit input and four hidden layers with $256, 128, 32$, and 16 units, respectively, a two-unit softmax output layer; and a training epoch count set to 5000, ADAM optimizer employing a learning rate equal to 0.001, a log-loss function. For every above-tested machine learning method, 10% of training data was applied for validation in a ten-fold cross-validation run, respectively.

III. RESULTS

Results of the behavioral (touchpad preference responses) experiment we summarized in Figure 1 for reminiscent (childhood age) versus modern/contemporary interiors within Western and Japanese designs. The response distributions from all 15 participating elderly persons have shown significantly lower preferences $(p_r(N* \;\; vs \;\; O*) \leq 0.05/4 = 0.0125$ as tested with Bonferroni-corrected Wilcoxon rank-sums tests) for the the modern Japanese (NJ) and Western (NW) designs. Furthermore, the experiment resulted in reminiscent Japanese interiors to be more favorable. The reminiscent Western interiors resulted in between preference results with also statistically significant differences to the other groups (see details in Figure 1). We also analyzed the subject preference distributions grouped for all available scores in the project. Higher MoCA scoring subjects resulted in unimodal while the lower scoring participants in bimodal distributions as shown in Figure 2. Finally, the subject interior evaluation preference scores from Figure 2 we also grouped according to normal (MoCA$>$ 25) and MCI (MoCA$\leq$ 25) scores. The result confirmed a previous observation of the MCI subjects' bimodal distributions with $p_k = 0.0575$ as obtained from the Kolmogorov-Smirnov test for distribution comparisons.

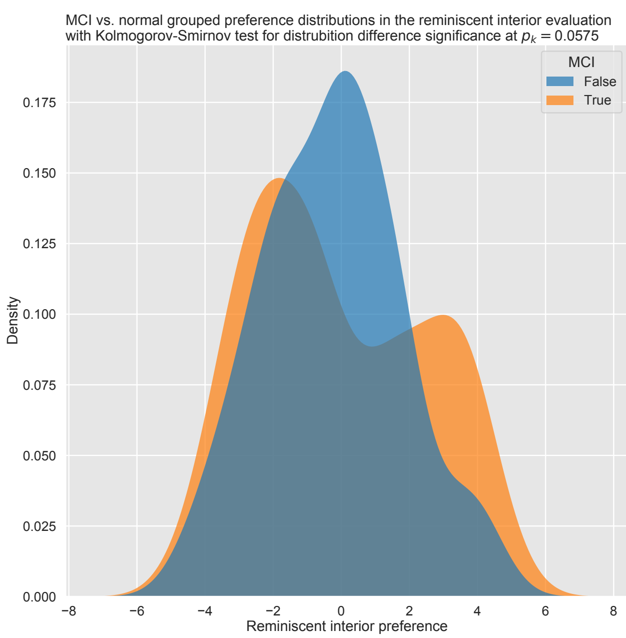

Fig. 3. The subject interior evaluation preference scores from Figure 2 grouped into normal and MCI (MoCA$\leq$ 25) levels. The MCI subjects results further confirmed the observation of a bimodal distribution with $p_k = 0.0575$ as obtained from Kolmogorov-Smirnov test for distribution comparisons.

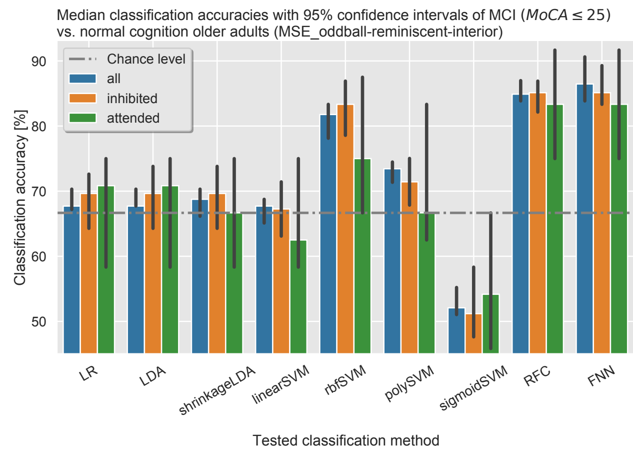

Fig. 4. Classification results of normal versus MCI subjects using EEG-derived MSE features from four EEG electrodes. A chance level in the experiment was of 66.7%. The best results were obtained for all oddball experiment response groups (attended, inhibited, or all together) for the fully connected neural network (FNN) and random forest (RFC) classifiers.

Results of EEG binary classification of normal (MoCA$>$ 25) versus MCI (MoCA$\leq$ 25) we summarized from bar-plots with error bars depicting 95% confidence intervals of classification results using the evaluated classifiers and three types of analyzed response groups as shown in Figure 4. The RFC and FNN classifiers resulted in the best median accuracies. All three groups attended (oddball targets), ignored (oddball non-targets), and all together EEG response classification were above 80% (a chance level was of 66% for our subject group) for the FNN classifier.

IV. CONCLUSIONS

The described project delivered two notable outcomes addressing behavioral and brain response distributions in a reminiscent interior evaluation task. We showed elderly subject preferences favoring reminiscent interior designs with childhood-age and own culture (here Japanese) environments, which confirmed our hypothesis of this kind of imagery usefulness for dementia biomarker task design. We also observed unimodal response distributions for normal versus bimodal (more extreme) for the MCI group based on MoCA evaluations. Furthermore, the binary classification framework of MCI (MoCA $\leqslant 25$) versus normal cognition (MoCA > 25) in a working oddball memory task using reminiscent interior images produced remarkable accuracies using MSE features derived from four EEG channels (median classification accuracies above 80% for the best methods using RFC and FNN) as outlined in Figure 4.

The AI/ML-based dementia onset forecasting successful adoption shall lead to healthcare cost-saving, benefiting aging societies globally. We additionally recognize the inherent limitations of the presented method. However, at this point, we reproduce human-error-prone and subjective cognitive evaluation measures, which are proxy prognostications of dementia origin. Furthermore, the current project involved a small sample of subjects, which is an additional limitation of the reported classification results.

Author Contributions

TMR: Conceived the concept of the reminiscent interior-based oddball experiment utilization for EEG and MSE application to MCI prediction with machine learning methods; TMR and MSA: designed and conducted EEG experiments with the elderly; MOM supported recruitment and management of participants; TMR created and programmed the data analysis; TMR, MSA, and TK: analyzed the data; TMR, MSA, MOM, and TK discussed results; TMR: wrote the paper.

ACKNOWLEDGMENTS

The authors thank all the elderly experimental participants from Tokyo Silver Human Resources Center who shared their time and participated in our study. The authors acknowledge discussion and feedback during the project initial stages by Dea Luma and Yoshiyuki Kawazoe from The University of Tokyo, Japan. This research was supported in part by the KAKENHI, the Japan Society for the Promotion of Science Grant No. JP18K18140 (MSA), JP18KT0035 (MOM), JP19H01138 (MOM), JP20H05022 (MOM), JP20H05574 (MOM), and the Japan Science and Technology Agency AIP Trilateral AI Research Grant No. JPMJCR20G1 (MOM, TMR).

References

[1] World Health Organization, "Risk reduction of cognitive decline and dementia: Who guidelines," in *Risk reduction of cognitive decline and dementia: WHO guidelines*, 2019.

[2] G. Livingston, J. Huntley, A. Sommerlad, D. Ames, C. Ballard, S. Banerjee, C. Brayne, A. Burns, J. Cohen-Mansfield, C. Cooper, S. G. Costafreda, A. Dias, N. Fox, L. N. Gitlin, R. Howard, H. C. Kales, M. Kivimäki, E. B. Larson, A. Ogunniyi, V. Orgeta, K. Ritchie, K. Rockwood, E. L. Sampson, Q. Samus, L. S. Schneider, G. Selbæk, L. Teri, and N. Mukadam, "Dementia prevention, intervention, and care: 2020 report of the Lancet Commission," *Lancet*, vol. 0, jul 2020.

[3] S. A. Sikkes, Y. Tang, R. J. Jutten, L. M. Wesselman, L. S. Turkstra, H. Brodaty, L. Clare, E. Cassidy-Eagle, K. L. Cox, G. Chételat, S. Dautricourt, K. Dhana, H. Dodge, R. M. Dröes, B. M. Hampstead, T. Holland, A. Lampit, K. Laver, A. Lutz, N. T. Lautenschlager, S. M. McCurry, F. J. Meiland, M. C. Morris, K. D. Mueller, R. Peters, G. Ridel, A. Spector, J. T. van der Steen, J. Tamplin, Z. Thompson, and A. Bahar-Fuchs, "Toward a theory-based specification of non-pharmacological treatments in aging and dementia: Focused reviews and methodological recommendations," *Alzheimer's Dement.*, no. May, pp. 1–16, 2020.

[4] P. Julayanont, Z. Nasreddine, M. Brousseau, M. Borrie, H. Chertkow, and N. Phillips, "The Montreal cognitive assessment memory index score (MOCA-MIS) and total MOCA score to help predict mci conversion to Alzheimer's disease," *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, vol. 8, no. 4, p. P372, 2012.

[5] M. Nara, M. Sugie, T. Takahashi, T. Koyama, R. Sengoku, Y. Fujiwara, O. Shuichi, K. Harada, S. Kyo, and H. Ito, "Japanese version of the montreal cognitive assessment cut-off score to clarify improvement of mild cognitive impairment after exercise training in community-dwelling older adults: Moca-j score to clarify mci improvement," *Geriatrics & Gerontology International*, vol. 18, 02 2018.

[6] H. Sugimoto, T. Kawagoe, and M. Otake-Matsuura, "Characteristics of resting-state functional connectivity in older adults after the picmor intervention program: a preliminary report," *BMC geriatrics*, vol. 20, no. 1, pp. 1–10, 2020.

[7] T. M. Rutkowski, Q. Zhao, M. S. Abe, and M. Otake, "AI neurotechnology for aging societies - task-load and dementia EEG digital biomarker development using information geometry machine learning methods," in *AI for Social Good Workshop at the Neural Information Processing Systems (NeurIPS) 2018*, (Montreal, Canada), pp. 1–4, December 8, 2018.

[8] T. M. Rutkowski, Q. Zhao, M. S. Abe, and M. Otake-Matsuura, "Passive BCI for task-load and dementia biomarker elucidation," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Berlin, Germany), p. ThC01.1, IEEE Engineering in Medicine and Biology Society, IEEE Press, July 23–27, 2019.

[9] T. M. Rutkowski, M. Koculak, M. S. Abe, and M. Otake-Matsuura, "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, May 2019.

[10] T. M. Rutkowski, M. S. Abe, M. Koculak, and M. Otake-Matsuura, "Cognitive assessment estimation from behavioral responses in emotional faces evaluation task - ai regression approach for dementia onset prediction in aging societies -," in *NeurIPS Joint Workshop on AI for Social Good*, vol. Track 1 - Producing Good Outcomes, (Vancouver, Canada), pp. 1–4, December 14, 2019.

[11] T. M. Rutkowski, M. S. Abe, M. Koculak, and M. Otake-Matsuura, "Classifying mild cognitive impairment from behavioral responses in emotional arousal and valence evaluation task - ai approach for early dementia biomarker in aging societies -," in *The 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Montreal, Canada), pp. 5537–5543, IEEE Engineering in Medicine and Biology Society, IEEE Press, July 20–24, 2020.

[12] O. E. Krigolson, C. C. Williams, A. Norton, C. D. Hassall, and F. L. Colino, "Choosing muse: Validation of a low-cost, portable eeg system for erp research," *Frontiers in Neuroscience*, vol. 11, p. 109, 2017.

[13] A. Barachant, D. Morrison, H. Banville, J. Kowaleski, U. Shaked, S. Chevallier, and J. J. T. Tresols, "muse-lsl." ZENODO, May 2019.

[14] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals," *Physical review E*, vol. 71, no. 2, p. 021906, 2005.

[15] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, "Analysis of electroencephalograms in alzheimer's disease patients with multiscale entropy," *Physiological measurement*, vol. 27, no. 11, p. 1091, 2006.

[16] Cycling '74, "Max 8," 2019.

[17] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.

[18] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, 2000.

[19] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. H. A. Chen, "Neurokit2: A Python toolbox for neurophysiological signal processing," *Behavior Research Methods*, Feb 2021.

[20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.