Soundscape Neuroarts and Cognitive Well-being: EEG and fNIRS Approach

Simon Kojima (simon.kojima@ieee.org)
RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan
Reiko Shiba (shiba.reiko@pc.geidai.ac.jp)
Tokyo University of The Arts, Tokyo, Japan
Yota Morimoto (yota@waseda.jp)
Waseda University, Tokyo, Japan
Takayuki Hamano (hamano.takayuki@kunitachi.ac.jp)
Kunitachi College of Music, Tokyo, Japan
Kiyoshi Furukawa (furukawa@fa.geidai.ac.jp)
Tokyo University of The Arts, Tokyo, Japan
Shin'ichiro Kanoh (kanoh@shibaura-it.ac.jp)
Shibaura Institute of Technology, Tokyo, Japan
Tomasz Komendziński (tkomen@umk.pl)
Nicolaus Copernicus University, Toruń, Poland
Mihoko Otake-Matsuura (mihoko.otake@riken.jp)
RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan
Tomasz M. Rutkowski (tomasz.rutkowski@riken.jp)
RIKEN Center for Advanced Intelligence Project (AIP), Tokyo, Japan

Abstract

Early detection of cognitive decline is crucial for maintaining well-being in aging populations. As a pilot study, this research explores the potential of auditory-evoked neurophysiological responses to predict cognitive function in young and middle-aged adults. We developed a deep learning model to estimate reaction time, a proxy for cognitive performance, from electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) data. Participants engaged in personalized soundscape creation and listening tasks, designed to elicit auditory processing relevant to communication. Our model demonstrated significant predictive accuracy, achieving an R^2 value exceeding 0.8 for attention-related cognitive assessments. Our findings suggest that the synergy of passive BCI cognitive assessment, conducted within active auditory paradigms, and deep learning analysis of neurophysiological data, represents a highly promising non-invasive strategy for objective cognitive monitoring. This methodological advancement offers considerable potential for the scalable early detection of cognitive decline, which could facilitate more timely therapeutic interventions and ultimately promote overall well-being.

Keywords: cognitive well-being; EEG; fNIRS; deep learning; neuroarts

Introduction

Maintaining cognitive health is paramount for well-being, particularly in aging populations. Cognitive functions, vital for daily living, are significantly compromised by conditions like dementia and mild cognitive impairment (Livingston et al., 2024). Recent research highlights the potential for dementia risk reduction, with interventions during middle age showing substantial promise (Livingston et al., 2024). Notably, addressing hearing loss, a key midlife risk factor, can significantly impact dementia incidence (Livingston et al., 2024). Given the critical role of hearing in communication and its influence on art experiences, as explored in the emerging field of Neuroarts, we hypothesized that auditory-evoked neural activity could serve as a reliable indicator of cognitive function. This pilot study investigated the feasibility of using EEG and fNIRS during personalized soundscape tasks to predict cognitive performance in young and middle-aged adults. A deep learning model was trained to estimate reaction times from neurophysiological data, achieving a mean absolute percentage error (MAPE) of 5.84%. These findings suggest that auditory-evoked neural responses during artistic engagement can effectively assess cognitive function, paving the way for passive brain-computer interfaces (BCIs) and digital neurobiomarkers for early detection of cognitive decline.

Methods

Five healthy participants, ranging in age from their twenties to their sixties, created personalized soundscapes using $Soundtope^{TM}$ for well-being (Shiba et al., 2023), a tablet-based application enabling music and environmental sound selection, volume balancing, and sound source localization. Audio was delivered via headphones (Sennheiser HD 280 Pro). Each pilot study participant completed two sessions (morning and evening) separated by their daily work schedule. Each session comprised four three-minute soundscape creation runs: music selection, environmental sound selection, mixing, and sound source localization. Following these, participants completed Task-Switching and Memorability cognitive tests (Adolphe et al., 2022), and reaction time (RT) was recorded. Concurrent with soundscape

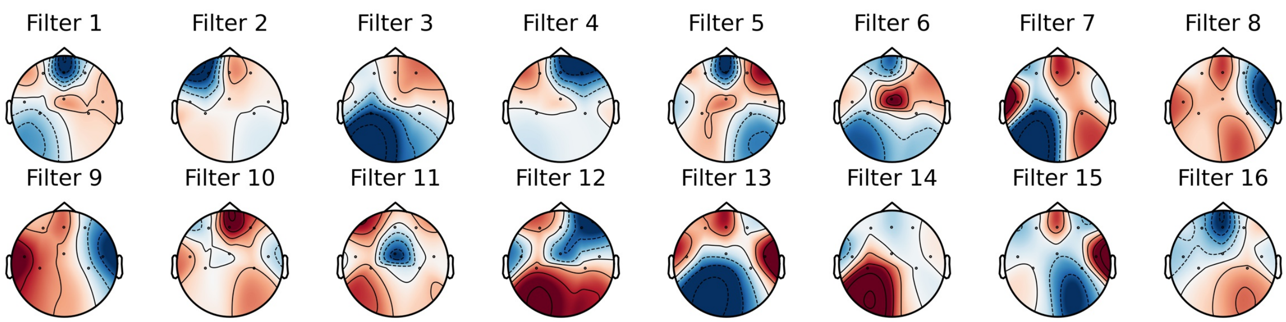

Figure 1: Spatial filters (kernel weights of the depthwise convolution layer) trained for EEG channels.

creation, eight-channel electroencephalography (EEG) and eight-channel functional near-infrared spectroscopy (fNIRS) data were acquired at 250 Hz using g.tec's Unicorn Hybrid Black and g.SENSOR fNIRS systems, respectively, for electrophysiological and neurovascular monitoring. This study was conducted in accordance with the Declaration of Helsinki and approved by the RIKEN Ethical Review of Research Involving Human Subject (Wako-shi, Saitama, Japan; Wako3 30-28(4)). Recorded EEG ($1 \sim 40$ Hz) and fNIRS ($0.1 \sim 0.4$ Hz) signals were bandpass filtered and segmented into nine $20-$second intervals per run. Reaction times (RTs) from Task-Switching and Memorability tests, averaged per session (morning/evening), served as regression targets. An EEGNet$-8,2$ model (Lawhern et al., 2018) was employed for regression, with modifications: the first Conv2D kernel size was $(1,125)$, the first AveragePool2D kernel size was $(1,8)$, and dropout was set to 0.5. The model output was two, corresponding to the RTs of the two cognitive tasks. Training used mean squared error (MSE) loss, the AdamW optimizer, a $500-$epoch maximum, and a batch size of 8. Ten-fold cross-validation was used, with 20% of the training data allocated for validation. The model achieving the lowest validation loss was selected for performance evaluation.

Results and Discussion

The prediction accuracies, presented in Table 1, demonstrate acceptable performance, with Mean Absolute Percentage Error (MAPE) values below 10% for both cognitive tests. Notably, the coefficient of determination (R^2) exceeded 0.8 for the Task-Switching Test, indicating a strong predictive capability of our methodology for this specific cognitive assessment.

Table 1: Prediction accuracies. Each abbreviation is as follows: MSE (mean squared error), MAE (mean absolute error), and MAPE (mean absolute percentage error). The value after $\pm$ denotes standard deviation across folds.

Index	Memorability	Task Switching
R^2	-0.32 ± 2.45	0.81 ± 0.11
MSE	4284 ± 8184	7814.0 ± 5781
MAE	31.18 ± 8.21	61.68 ± 9.92
Median Error	24.29 ± 8.80	50.01 ± 11.10
MAPE (%)	3.68 ± 1.02	8.01 ± 1.48

Conversely, the negative R^2 value observed for the Memorability Test suggests that our model performed worse than a simple mean prediction. However, the substantial standard deviation across cross-validation folds implies significant variability in model performance, potentially indicating suboptimal training. Future work should focus on optimizing the model architecture, refining learning parameters, and increasing the sample size to improve predictive accuracy for this task.

Figure 1 displays the spatial filters (depthwise convolution kernel weights) for EEG channels, obtained from the fold exhibiting the highest prediction accuracy. Due to the limited EEG channel coverage (CP line and forward, according to the $10-20$ system), the topographic map in the parietal-occipital region was extrapolated. The spatial filters reveal localized feature extraction, exemplified by filters 6 and 11 focusing on EEG channel Cz. This suggests that informative features for cognitive performance estimation may be sparsely distributed across frequency and channel domains. Consequently, it may be feasible to achieve comparable prediction accuracy with a reduced channel count and fewer temporal convolution filters, facilitating the development of practical applications.

Conclusions

This study demonstrates the feasibility of predicting reaction times in cognitive tests using neurophysiological signals acquired during personalized soundscape creation and listening tasks, a Neuroarts-based approach. Specifically, we have shown that auditory engagement within a creative context elicits neural responses that correlate with cognitive performance.

These findings suggest the potential for developing a non-invasive, passive BCI neurophysiological assessment tool for early detection of cognitive decline. By integrating this methodology into a practical system, we aim to facilitate timely interventions and personalized treatments for individuals at risk of dementia and mild cognitive impairment. This approach could contribute to a scalable and accessible solution for monitoring cognitive health, ultimately promoting well-being within aging populations. This work lays the foundation for future research integrating Neuroarts paradigms with advanced machine learning for early disease detection and personalized cognitive health.

References

Adolphe, M., Sawayama, M., Maurel, D., Delmas, A., Oudeyer, P.-Y., & Sauzéon, H. (2022, June). An Open-Source Cognitive Test Battery to Assess Human Attention and Memory. *Frontiers in Psychology*, *13*. doi: 10.3389/fpsyg.2022.880375

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018, July). EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, *15*(5), 056013. doi: 10.1088/1741-2552/aace8c

Livingston, G., Huntley, J., Liu, K. Y., Costafreda, S. G., Selbæk, G., Alladi, S., . . . Mukadam, N. (2024, August). Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *The Lancet*, *404*(10452), 572–628. doi: 10.1016/S0140-6736(24)01296-0

Shiba, R., Morimoto, Y., Akimoto, R., Nageishi, K., Akechi, T., Fujimori, M., . . . Furukawa, K. (2023, December). Verification of the effect of a music application for the purpose of developing medical applications-. *IEICE Technical Report Report*, *HIP2023-82*, 31-35.