DOI: 10.3217/978-3-99161-050-2-062

Hardware Friendly Corticomorphic Hybrid CNN-SNN Architecture for EEG-Based Auditory Attention Detection

Richard Gall[1*], Deniz Kocanaogullari[1], Murat Akcakaya[1], Deniz Erdogmus[2], Rajkumar Kubendran[1]
[1]*University of Pittsburgh, Pittsburgh, PA, USA;* [2]*Northeastern University, Boston, MA, USA*
E-mail: rtg16@pitt.edu

Introduction: For a healthy listener, the ability to selectively attend to a particular speaker in a multiple-speaker, or ”cocktail party” [1], scenario is a trivial task. Those who suffer from hearing loss or are hearing impaired have difficulty with selective auditory attention [2]. Cognition and selective auditory attention of an individual play an important role in listening and communication [2], but unfortunately, these are currently not fully utilized during hearing aid design [3].
Studies into the auditory attention network of the brain have shown that it is possible to decode auditory attention through the neural responses of a listener, suggesting that assisted selective auditory attention is possible, which led to developments in auditory attention decoding (AAD) using EEG-based BCIs. The current AAD methods are impractical for hearing aids due to large model architectures, requiring many operations, that when ported to hardware call for a sizable memory footprint, more compute resources, and higher power consumption than realistic.

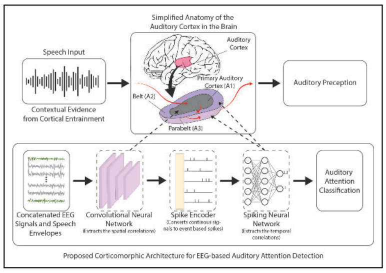

Figure 1: *The proposed corticomorphic hybrid architecture, inspired by the anatomy of the auditory cortex. Speech input and contextual evidence from other regions in the brain are passed into the primary auditory cortex (A1). The processed signals are then transmitted to the belt (A2), and finally the parabelt (A3). We hypothesize that the convolution layer will mimic the primary auditory cortex (A1) while the spiking neural network layer will emulate the belt (A2) and parabelt (A3).*

Material, Methods and Results: Our work demonstrates a pioneering approach to developing a hardware-friendly corticomorphic neural network modeled to mimic the layered cortical structure of the brain. Our proposed hybrid network is biologically inspired by the organization of the auditory cortex, Fig. 1. The proposed hybrid convolutional and spiking neural network (CNN-SNN) is evaluated on our own EEG-based BCI dataset [4] (10 participants) and, additionally, compared to the current state-of-the-art AAD methods on a publicly available dataset, DTU [5] (18 participants). The CNN-SNN achieved an accuracy of 98.8% on our dataset with int8 precision, while utilizing 50% fewer EEG channels with an $\sim 5\%$ decrease in operations. Similarly, the model was able to outperform all state-of-the-art models on the DTU dataset by $\sim 10\%$ for the 0.5, and 1 second decision windows, while using 87.5% fewer EEG channels and int8 precision. Hence, the overall benefits of using fewer channels, reducing the memory footprint with a lower bit precision, and a shorter decision window.

Conclusion: The proposed model architecture has surpassed current state-of-the-art results for AAD. The CNN-SNN architecture performs within a more desirable decision window (0.5 or 1 sec), uses fewer EEG channels, and increases the accuracy compared to state-of-the-art models. These results indicate that with the proposed model, it is possible to accurately decode auditory attention with high accuracy and less power consumption and smaller memory footprints compared to a conventional ANN.

Acknowledgments and Disclosures: There are no conflicts of interest to disclose.

References:

[1] E. C. Cherry, “Some experiments on the recognition of speech, with one and with two ears,” *The Journal of the acoustical society of America*, vol. 25, no. 5, pp. 975–979, 1953.

[2] B. G. Shinn-Cunningham and V. Best, “Selective attention in normal and impaired hearing,” *Trends in amplification*, vol. 12, no. 4, pp. 283–299, 2008.

[3] B. Taylor and D. Hayes, “Does current hearing aid technology meet the needs of healthy aging,” *Hearing Review*, vol. 22, no. 2, pp. 22–26, 2015.

[4] M. Haghighi, M. Moghadamfalahi, M. Akcakaya, B. G. Shinn-Cunningham, and D. Erdogmus, “A graphical model for online auditory scene modulation using eeg evidence for attention,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1970–1977, 2017.

[5] S. A. Fuglsang, T. Dau, and J. Hjortkjær, “Noise-robust cortical tracking of attended speech in real-world acoustic scenes,” *Neuroimage*, vol. 156, pp. 435–444, 2017.

Published by Verlag der Technischen Universität Graz