Contribution ID: **41** Type: **Poster**

Sleep Spindle (LFADs) Project Abstract

Monday 10 July 2023 19:00 (2 hours)

A specific type of Electroencephalography (EEG) signals, sleep spindle, is believed to contribute to neuronal plasticity and memory consolidation. In this project, we proposed a system that is based on ultra-low latency and power FPGA to detect and interact with the sleep spindles to further understand the mechanism behind the theory. The proposed system will have a programmed FPGA that connects with a headstage. The headstage will record the subject's brain signals and the FPGA will process the signals to detect and interact with the sleep spindles.

Latent Factor Analysis via Dynamical Systems (LFADs) is the baseline deep learning model for this project. It is an RNN variational autoencoder for analyzing spiking neural data. LFADs follows the encoder-decoder structure. The input spiking data will be sent to a bidirectional GRU, Gaussian sampling, a unidirectional GRU, and several dense layers to produce the final outputs. LFADs will generate two vital outputs. One is a set of low-dimensional temporal factors which contains the information of the input spiking data. The other is the log firing rate, which can be translated to the firing rate that can generate the input spiking data.

We removed the Gaussian sampling from LFADs in this current project and will add the sampling layer back later. In this case, we have successfully deployed the no-sampling LFADs onto FPGA by implementing the unsupported layers in HLS4ML. The FPGA we used for the current deployment is Xilinx Alveo U50. By running LFADs on this board, we have significantly decreased the latency while maintaining reasonable model performance. We are currently working on optimizing the model by applying quantization aware training (QAT) on LFADs and managing to deploy multiple LFADs onto FPGA to enable high throughput processing. The high throughput processing can be used for analyzing largescale neural recordings and the low latency processing will allow neuroscientists to develop closed-loop technology to analyze the neural signal in real time.

Author: LIU, Xiaohan

Presenter: LIU, Xiaohan

Session Classification: Working dinner