A Convolutional Neural Network for Enhancing the Detection of SSVEP in the Presence of Competing Stimuli

Aravind Ravi, Jacob Manuel, Nargess Heydari and Ning Jiang*

27/07/2019

Engineering in Medicine and Biology Conference (EMBC 2019) Berlin, Germany

***Presented by: Dr. Ning Jiang**

eBionics Lab, Department of Systems Design Engineering

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Waterloo Engineering Bionics Lab

Outline

- Motivation and Proposed Method
- Methods – Stimulus and Data Acquisition
- Experiment
- Convolutional Neural Network (CNN)
- Canonical Correlation Analysis (CCA)
- Performance Evaluation
- Results
- Discussions
- Future Work

Motivation

Stimulus proximity has been shown to have an influence on the classification performance of a steady-state visual evoked potential based brain-computer interface (SSVEP-BCI)

Multiple visual stimuli placed close to each other compete for neural representations leading to the effect of competing stimuli[1]

This limits the range of flexibility for SSVEP stimulus interface design

[1]Kian B. Ng, Andrew P. Bradley, and Ross Cunnington. Stimulus specificity of a steady-state visual-evoked potential-based brain-computer interface. Journal of Neural Engineering, 2012

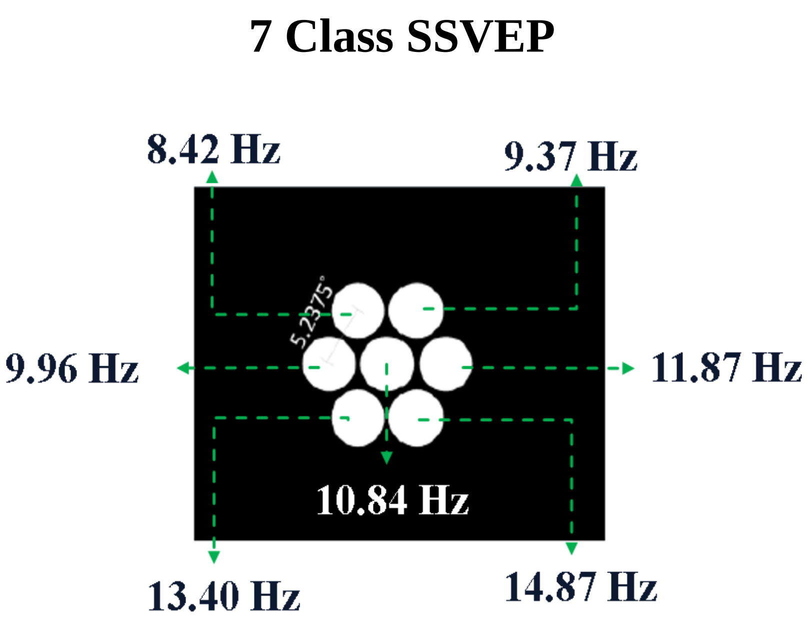
Proposed Method

A Convolutional neural network (CNN) based classification to enhance the detection accuracy of SSVEP in the presence of competing stimuli

A 7-class SSVEP dataset from 10 healthy participants was used for evaluating the performance

The results were compared with the canonical correlation analysis (CCA) detection algorithm

Methods – Stimulus and Data Acquisition



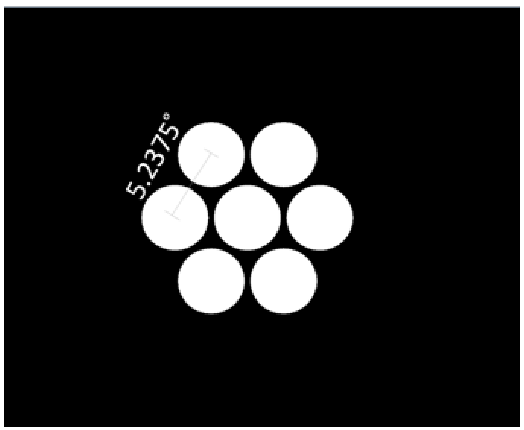
The g.USBamp and Gammabox (g.tec Guger Technologies, Austria) wet electrode (g.Scarabeo) system

Sampling rate - 1200 Hz
Channels - O1, O2, Oz, PO3, POz, PO4 and FPz
FPz – ground; right ear lobe – reference

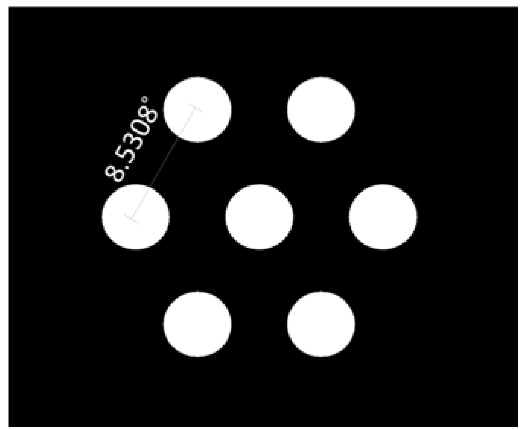
Only O1, O2 and Oz were used in the study

[2]Y. Wang, Y.-T. Wang, and T.-P. Jung. Visual stimulus design for high- rate SSVEP BCI. Electronics Letters, 46(15):1057, 2010.
[3]Masaki Nakanishi, Yijun Wang, Yu Te Wang, Yasue Mitsukura, and Tzyy Ping Jung. Generating visual flickers for eliciting robust steady- state visual evoked potentials at flexible frequencies using monitor refresh rate. PLoS ONE, 2014.

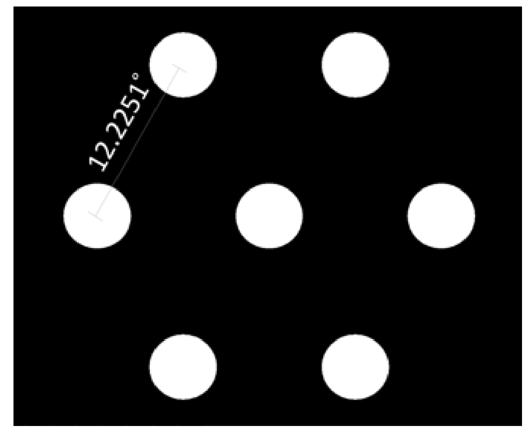
Stimulus Designs Evaluated



S1 – 5.24°



S2 – 8.53°



S3 – 12.23°

[3]A. Ravi, S. Pearce, X. Zhang, N. Jiang, and S. Member, “User-Specific Channel Selection Method to Improve SSVEP BCI Decoding Robustness Against Variable Inter-Stimulus Distance,” *9th Int. IEEE EMBS Conf. Neural Eng.*, pp. 283–286, 2019.

Experiment

12

12 s - Trial Length
2 s – Cue
6 s – Stimulation
4 s – Break

8

Total Number of Trials for each Stimulus

56

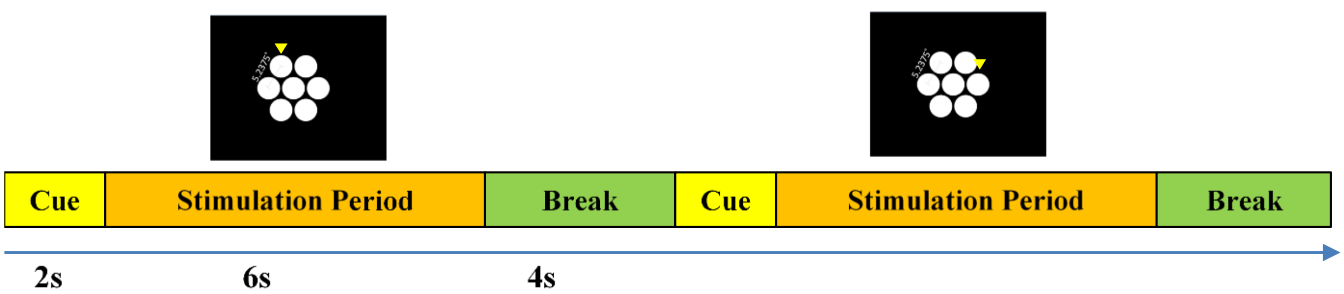
Total Number of Trials

3

Total number of Sessions
1 for each ISD

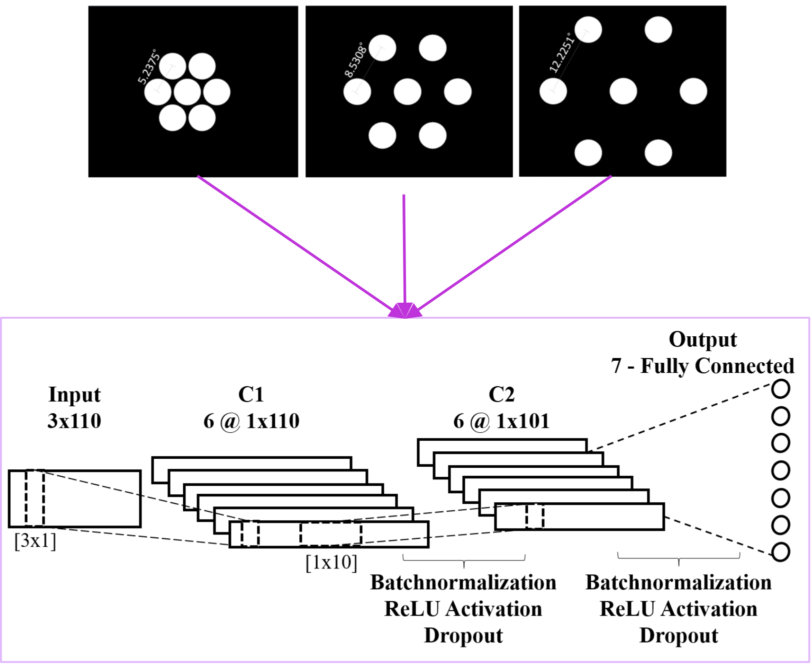
10

Total number of Participants



Convolutional Neural Network (CNN)

In comparison to the CNNs using time-domain inputs, a CNN using **frequency-domain inputs** would have a similar but **relatively simple network structure** and **reduced computational complexity (reduced number of tunable parameters)**



Investigated whether the **CNN parameters learned on one inter-stimulus distance (ISD) can generalize across to other ISDs and sessions**

[4]No Sang Kwak, Klaus Robert M¨uller, and Seong Whan Lee. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. PLoS ONE, 2017.

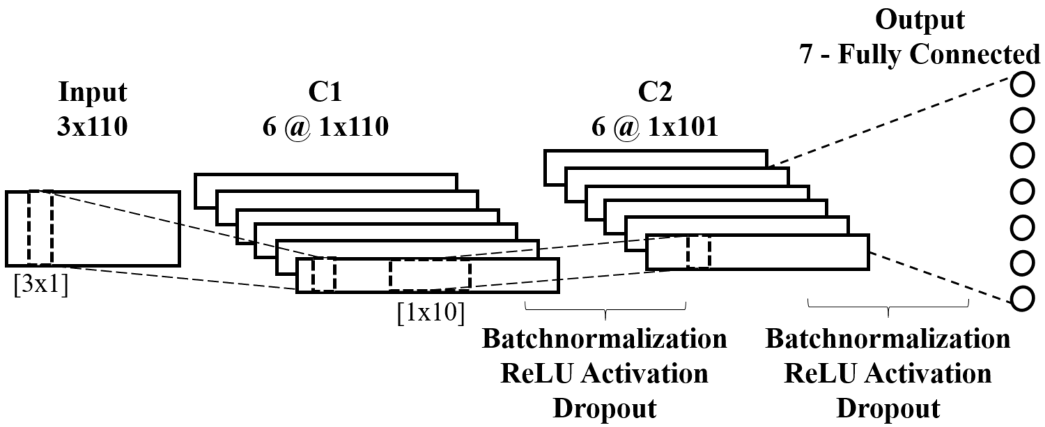
Convolutional Neural Network (CNN)

Pre-Processing

Filter: 4th order Butterworth band-pass filter between **1 Hz – 40 Hz**
Data Length: **1 s window sliding 100 ms**

FFT Features (Magnitude Spectrum)

Frequency Resolution: 0.2930 Hz
Frequency Components: 3 Hz to 35 Hz



[4]No Sang Kwak, Klaus Robert M¨uller, and Seong Whan Lee. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. PLoS ONE, 2017.

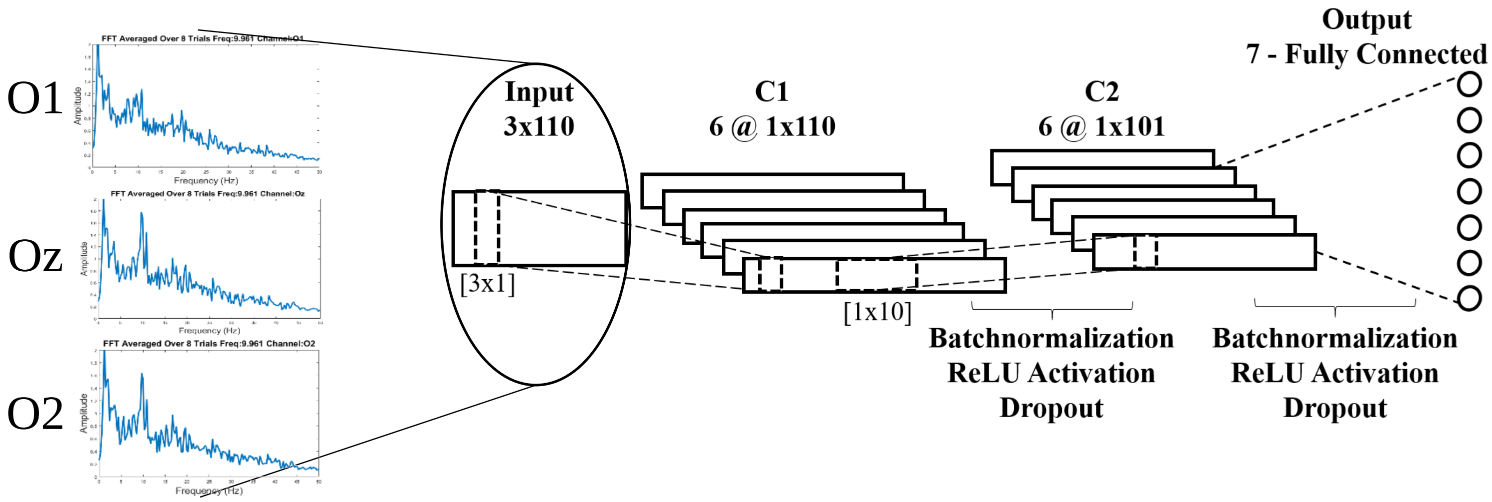
Convolutional Neural Network (CNN)

Pre-Processing

Filter: 4th order Butterworth band-pass filter between **1 Hz – 40 Hz**
Data Length: **1 s window sliding 100 ms**

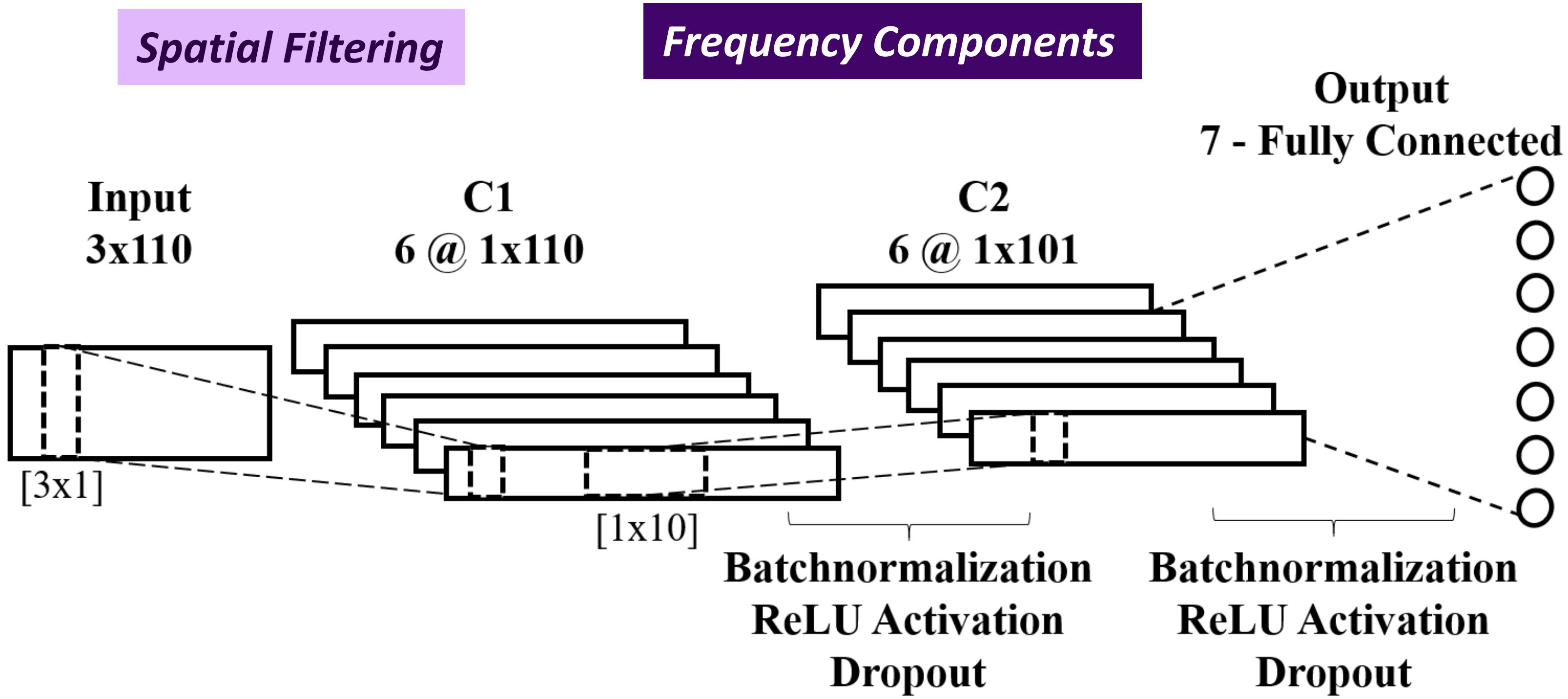
FFT Features (Magnitude Spectrum)

Frequency Resolution: 0.2930 Hz
Frequency Components: 3 Hz to 35 Hz



[4]No Sang Kwak, Klaus Robert M¨uller, and Seong Whan Lee. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. PLoS ONE, 2017.

Convolutional Neural Network (CNN)

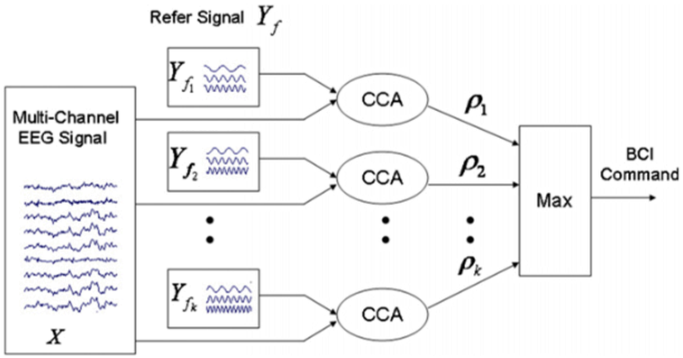


[4]No Sang Kwak, Klaus Robert M¨uller, and Seong Whan Lee. A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. PLoS ONE, 2017.

Canonical Correlation Analysis (CCA)

$$\rho(x,y) = max_{w_x,w_y} \frac{\mathrm{E}\left[w_x^T X Y^T w_y\right]}{\sqrt{\mathrm{E}[w_x^T X X^T w_x]\mathrm{E}\left[w_y^T Y Y^T w_y\right]}}$$

$$Y_n = \begin{bmatrix} sin(2\pi f_n t) \\ cos(2\pi f_n t) \\ \cdot \\ \cdot \\ sin(2\pi N_h f_n t) \\ cos(2\pi N_h f_n t) \end{bmatrix}, t = \left[\frac{1}{f_s}, \frac{2}{f_s}, \ldots, \frac{N_s}{f_s}\right],$$



The canonical correlation features ρ_{fi}, where i=1,2,...,7

C = argmax(ρ_{fi})

[5] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-Based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, 2007.
[6] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *J. Neural Eng.*, 2009.

Performance Evaluation

Offline Analysis

Simulated Online Analysis

<u>Training</u>

Hyper-parameters were optimized based on a **grid search**:

- Batch size (B): 2^b *, b* ∈ {5,6,7,8,9}
- Dropout Ratio (D): {0.25,0.3.0.35,0.4,0.45,0.5}
- Number of Epochs (E): {20,30,40,50,60}
- Learning Rate {α}: {0.001,0.002,0.005,0.01,0.1}
- Weights were initialized based on a Gaussian distribution (μ=0, σ^2=0.01)
- **Final Parameters were chosen** based on the **ones that generally resulted in the best performance for all participants**.

<u>Evaluation</u>

Case 1: CNN trained on S1 and tested on S2 and S3

Case 2: CNN trained on S2 and tested on S1 and S3

Case 3: CNN trained on S3 and tested on S1 and S2

Average accuracies across stimulus-distances were calculated

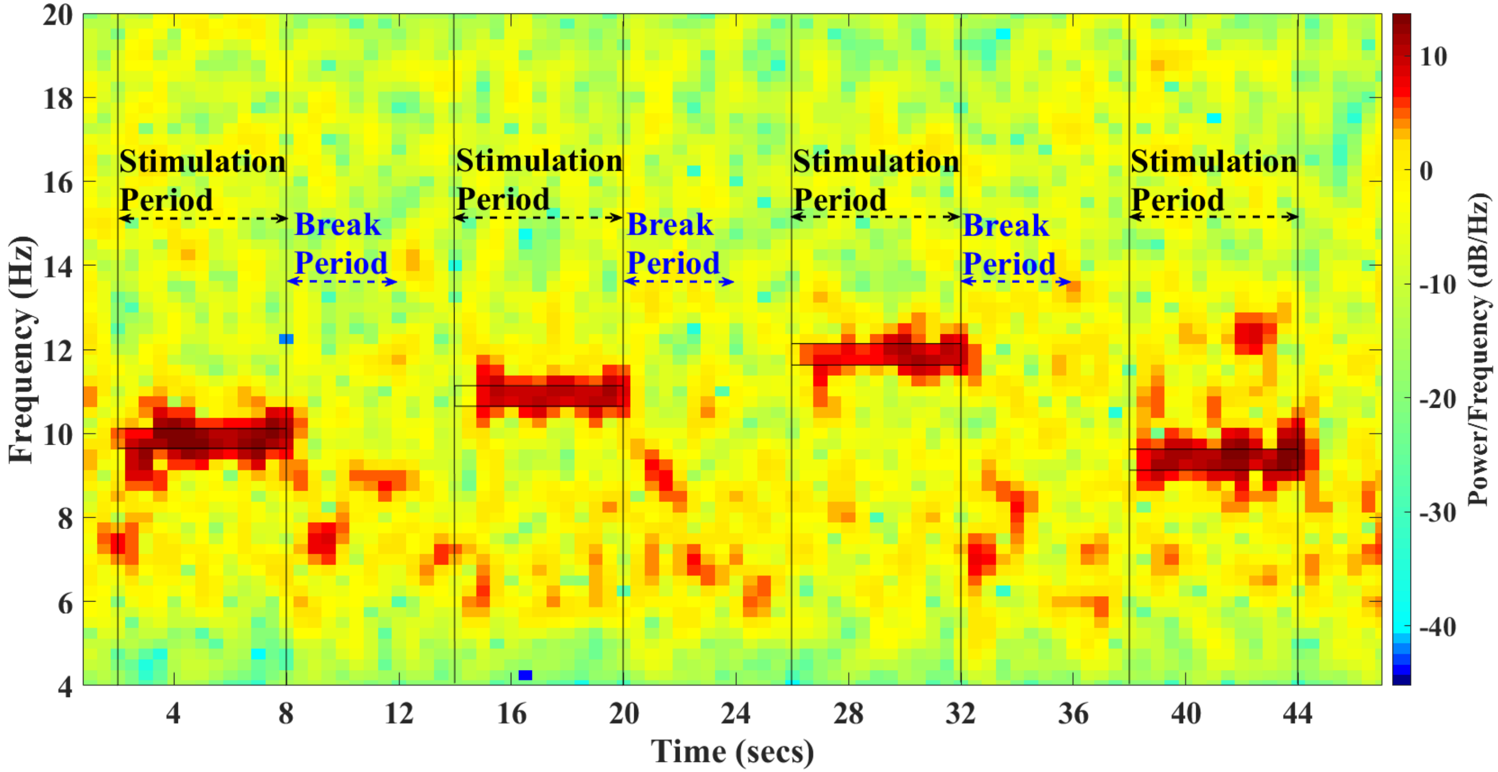
Performance Evaluation

Offline Analysis

Simulated Online Analysis

- All three cases as outlined previously were tested
- For the test data, the initial 1 s of each trial [0.5s 1.5s] from the start of the flickering period was considered, whereas the training data was segmented in the same manner as mentioned earlier
- The **classification accuracies** and the **information transfer rate (ITR)** were calculated for the CCA and the proposed CNN methods individually

Results

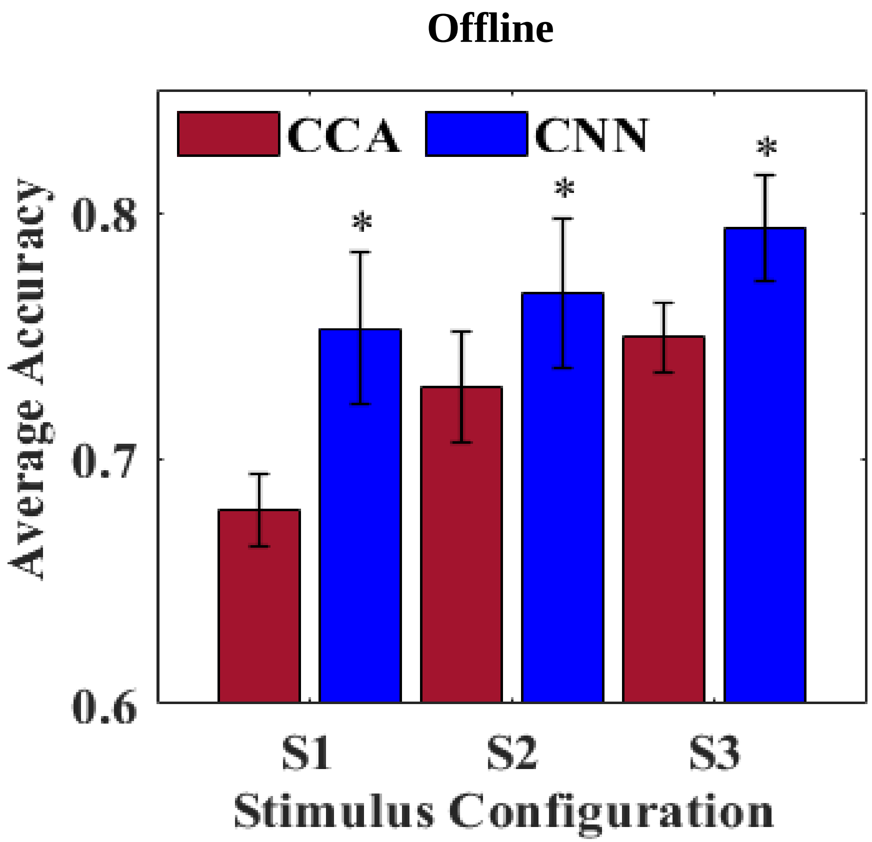


Illustrates an example of the spectrogram of four consecutive trials of SSVEP signals at frequencies 9.961 Hz, 10.84 Hz, 11.87 Hz and 9.375 Hz collected over the channel Oz

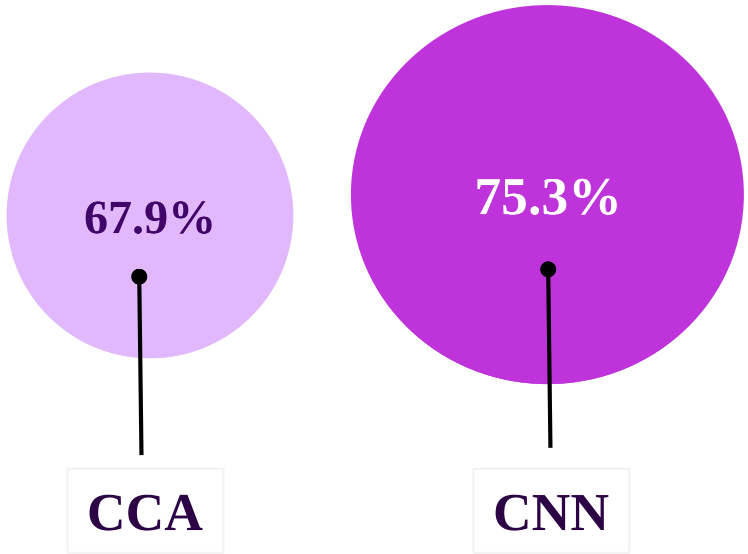
UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Results - Performance

S3 > S2 > S1 (p < 0.001)



Offline Analysis



S1 – Closest ISD (p < 0.001)

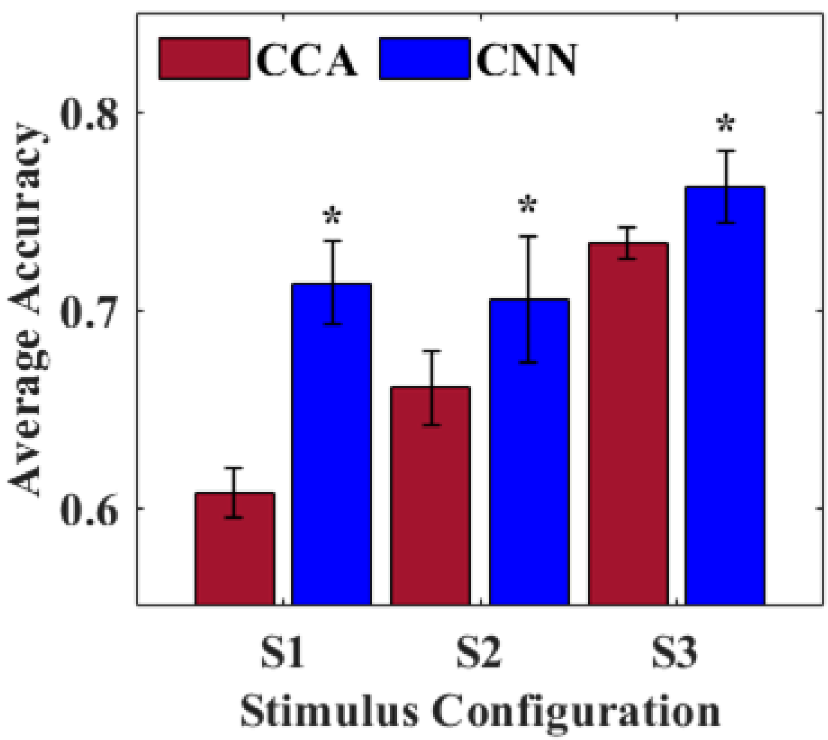
Results - Performance

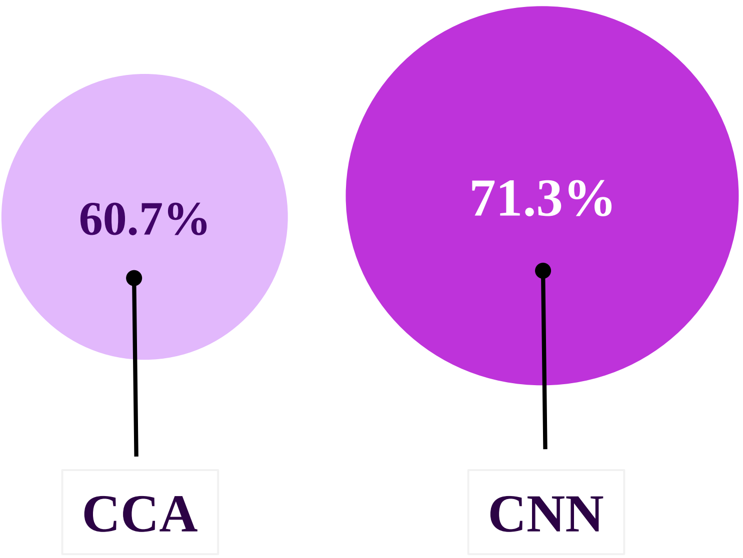
ITR (CNN vs. CCA) - (bits/min)

S1	S2	S3
51.0 vs. 34.4	51.3 vs. 42.6	59.0 vs. 52.5

Simulated Online Analysis

Simulated Online





S1 – Closest ISD ($p < 0.001$)

Results – Computational Complexity

The total number of trainable parameters were 4663

Trained on Intel Core i5-8400 CPU @ 2.80 GHz and 8 GB RAM

Overall training time was 6 s

Mean inference time for all segments was found to be 1.3 ms

Discussions

Average accuracy increased by over 10% using CNN on the closest ISD which is the most challenging case with the most significant completing stimuli

The CNN is robust in **decoding SSVEP across different ISDs**, and can be **trained independent of the ISD** resulting in a model that generalizes to other ISDs

Beneficial for practical applications developed on **virtual reality or augmented reality platforms** where the **stimuli would tend to be very closely spaced**

Interface Design - Provides more flexibility as **newly configured stimulus** distances can be **easily modified** with a **simple software update** and **retain the same CNN weights for inference**

Low computational load, short calibration time (approx. 12 minutes) and a 3-channel setup

Future Work (In Progress)

The proposed CNN across time windows and effects on performance when including more channels will be evaluated

Use of complex spectrum features and compare with other variants of CCA

User-independent training of a CNN

UNIVERSITY OF WATERLOO

FACULTY OF ENGINEERING

Waterloo Engineering Bionics Lab

Contact:
ning.jiang@uwaterloo.ca

eBionics Lab
https://uwaterloo.ca/engineering-bionics-lab/