Learnable Wavelet Neural Network for P300 Electroencephalogram Classification

Jason Stock and Chuck Anderson

Department of Computer Science, Colorado State University, 2021

Problem and Motivation

Accurate classification of electroencephalogram (EEG) signals is of paramount importance in brain-computer interface (BCI) systems - enabling communication without neuromuscular control.

Many existing works rely on time-frequency analysis with pre-defined filter-banks to decode EEG signals, but dominate frequencies may not be known prior, which can lead to inaccurate and uninterpretable results.

Scientific question: can we learn a wavelet filter-bank specialized to fit non-stationary signals and improve interpretability and performance for digital signal processing?

Datasets Details

EEG data recorded using the g.GAMMASys active electrodes (8 channel and 256Hz sampling rate) from the BCI research group at CSU are assessed for this study. We use a single motor impaired subject using a **Serial P300 Speller** to classify target stimuli.

All 120 target and 120 non-target samples are segmented for 812ms with a 78ms lead prior to the stimulus, baseline corrected, and bandpass filtered between 0.5 and 30Hz using a second order Butterworth filter.

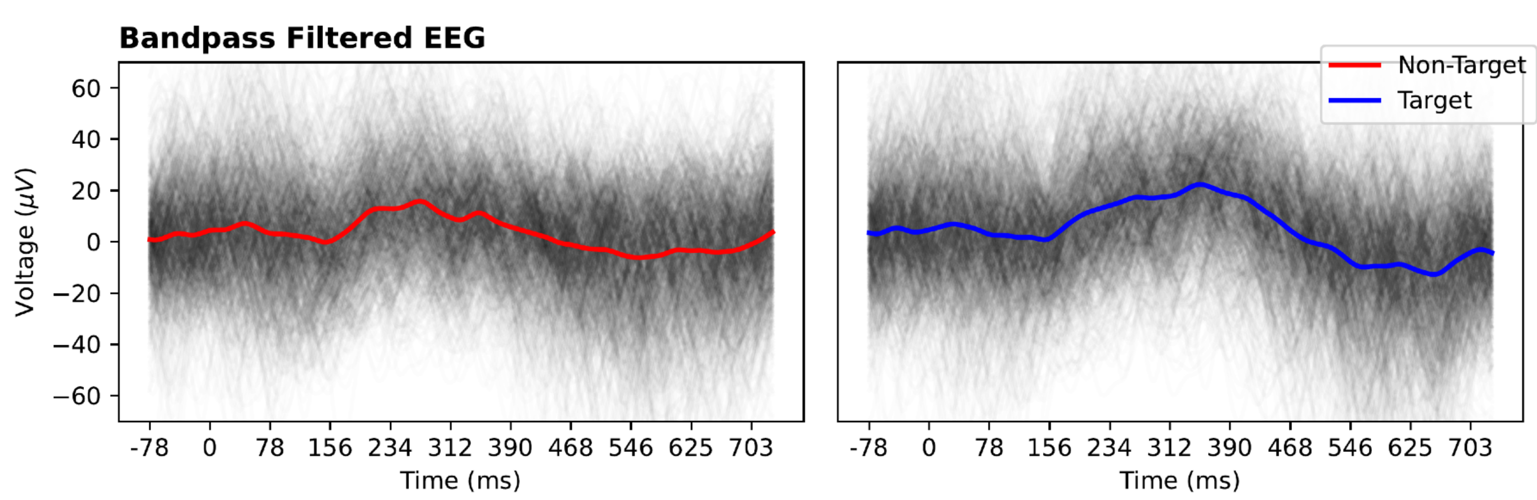

Figure 1: Class specific P300 signals with sample and channel mean.

Continuous Wavelet Transform

We can decompose our input signal, $x(t)$, to represent the time-frequency components (amplitude and phase) by convolving with a wavelet, $\psi(t)$, as:

$$CWT(s,\tau) = \frac{1}{|s|^{1/2}} \int x(t)\,\psi^*\left(\frac{t-\tau}{s}\right) dt\,.$$

For this work we use the **complex Morlet wavelet** as defined by:

$$\psi(t) = \underbrace{\left(s_t(2\pi)^{-\frac{1}{2}}\right)^{-\frac{1}{2}}}_{\text{Normalization}}\ \underbrace{\exp\left(\mathrm{i}2\pi f t\right)}_{\text{Complex sinusoid}}\ \underbrace{\exp\left(-\frac{t^2}{2s_t^2}\right)}_{\text{Gaussian envelope}}, \text{ where } s_f = \frac{f}{w},\ s_t = \frac{1}{2\pi s_f}\,.$$

The wavelet with frequency f and time t is normalized to have unit energy and truncated it on the interval $[-3.5s_t,\ 3.5s_t]$ to capture $99.98\,\%$ of the wavelet's energy.

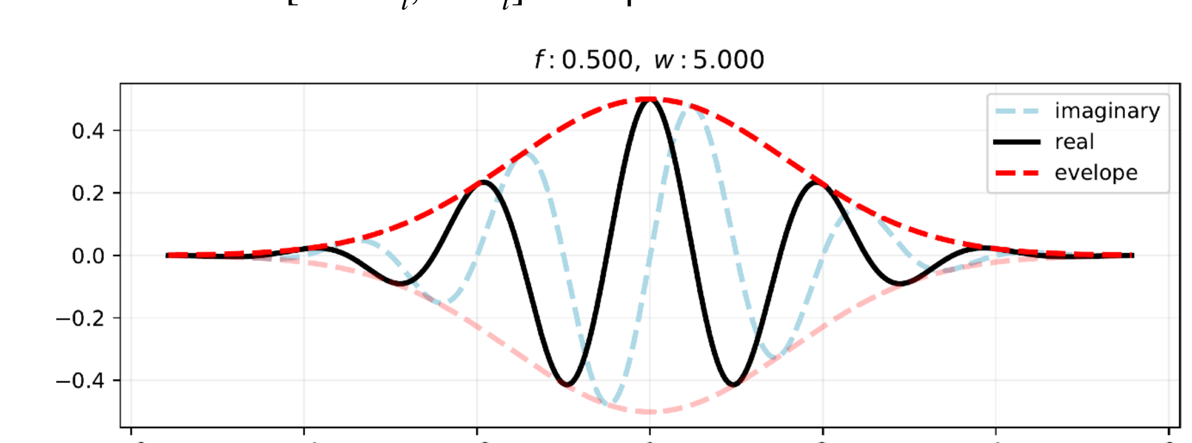

Figure 2: Complex Morlet Wavelet.

Training Details

Our proposal: use a CWT as the first layer of a neural network where the convolution is a parameterized function of a complex mother wavelet. Specifically, we use. $\psi_{f,w}(t)$ where frequencies, f, and widths, w, are trainable parameters learned through backpropagation.

The k-convolutional variable length filters (wavelets) are updated after every mini-batch to converge on frequencies that best represent the targets. Both parameter values f and w are clipped between $[0.5,\ 30]$Hz and $[4,\ 10]$, respectively.

Optimizer: Adam; Objective Function: Cross Entropy; with different learning rates, η:

$$(f,w)_t^0 = (f,w)_{t-1}^0 - \eta_0 \nabla_{(f,w)^0} LL(\theta),$$
$$\theta_t^{1:} = \theta_{t-1}^{1:} - \eta_1 \nabla_{\theta^{1:}} LL(\theta),$$
where $\eta_0 = 0.1,\ \eta_1 = 0.0001.$

Proposed Architecture

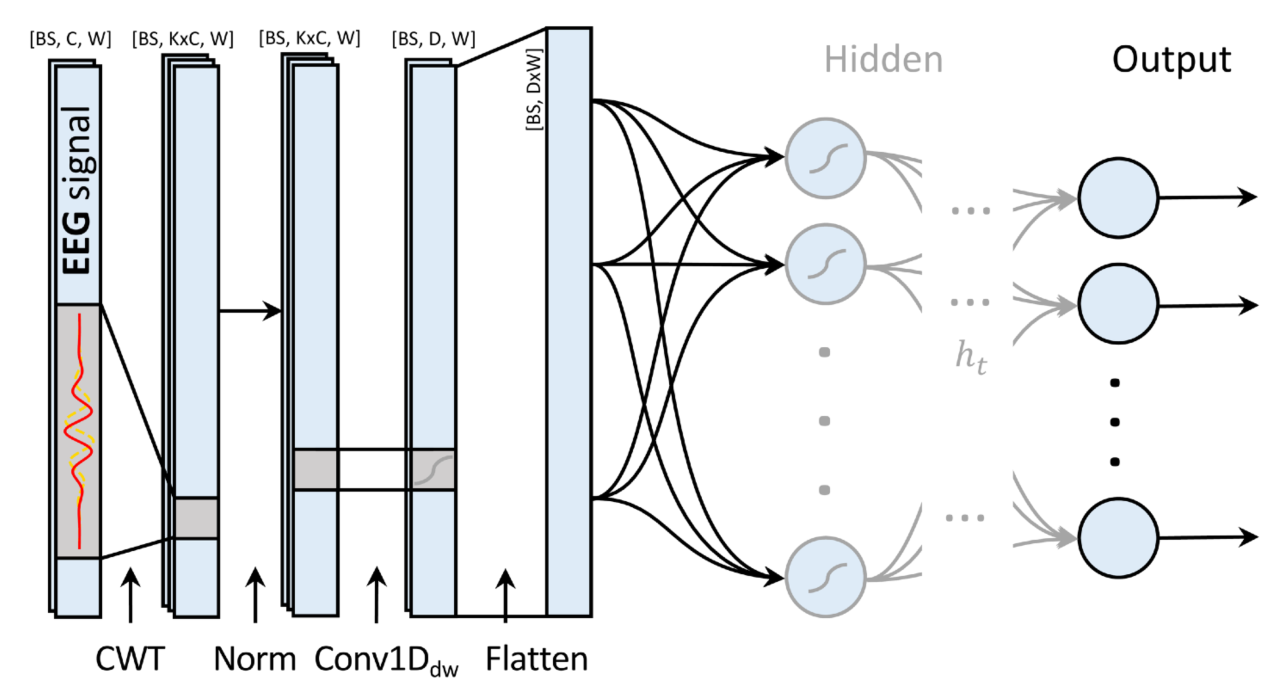

Figure 3: Network architecture.

Initial parameters: number of wavelets: $k = 32$; frequency range: $[0.5,\ 30]$ Hz; width range: $[5,\ 10]$; number of depth-wise convolution filters: $D = 3$.

Fully-connected hidden layers are optional and are not used in the present results. As such, the output from Conv1D$_{dw}$ goes to the linear output layer.

Results

We report accuracy of correctly predicted targets from our cwt-net and compare results to traditional approaches, including: fc-net, conv-net, and (filter)-bank-net.

Each model is trained and evaluated on the test data 10 times with different random weight initializations. The cwt-net yields a $81.25 \pm 3.48\,\%$ accuracy and **outperforms all other explored architectures**.

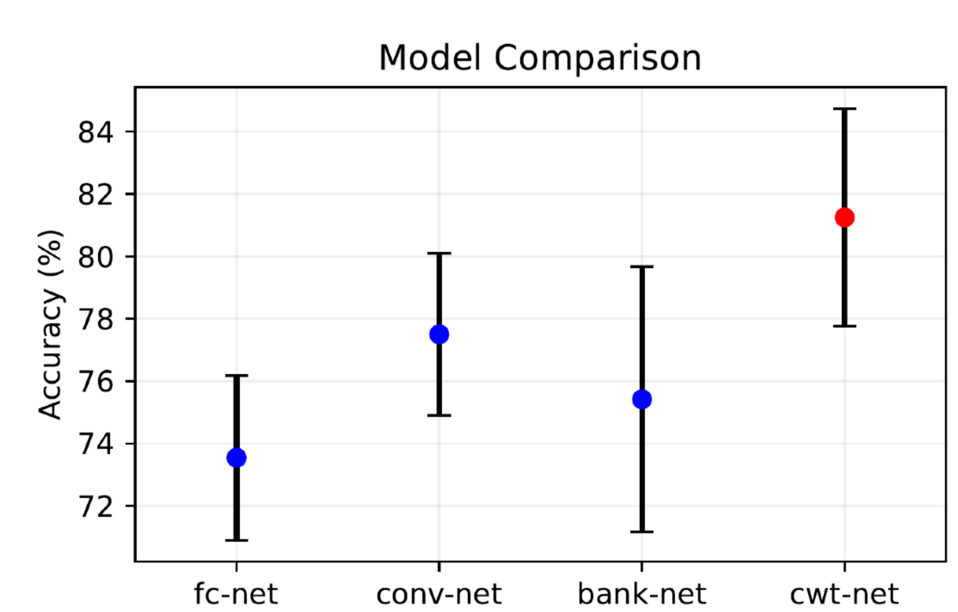

Figure 4: Mean and error in accuracy over multiple training trials.

Interpretations

Interpretations can be made after training the cwt-net to see what frequencies the network converges to and how these could potentially influence model performance.

In the forward pass, the real and imaginary convolutional output are combined as the complex transform and then the magnitude propagates to subsequent layers.

Figure 5 shows the magnitude scalogram with a high value at 10Hz through the onset of the stimulus and constant low frequencies following the P300.

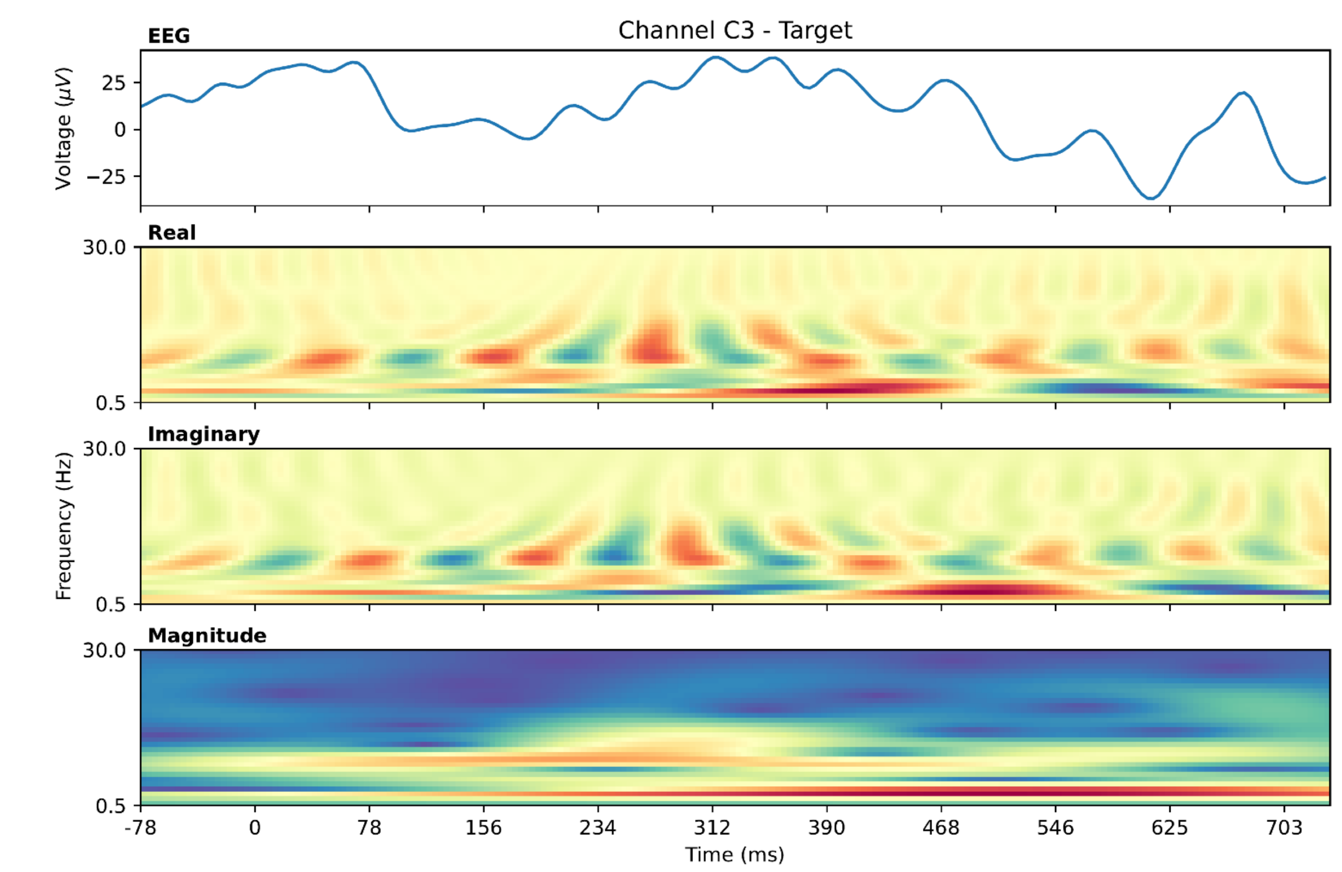

Figure 5: Example continuous wavelet transform over a single channel.

Multiple filters (wavelets) are linearly initialized and then updated during training. We find that multiple filters **converge on the same frequencies** (circled in Figure 6).

Filters with the same f_i and w_i are redundantly capturing the same features in the data. We speculate that reducing k can enable the network to learn unique filter parameters.

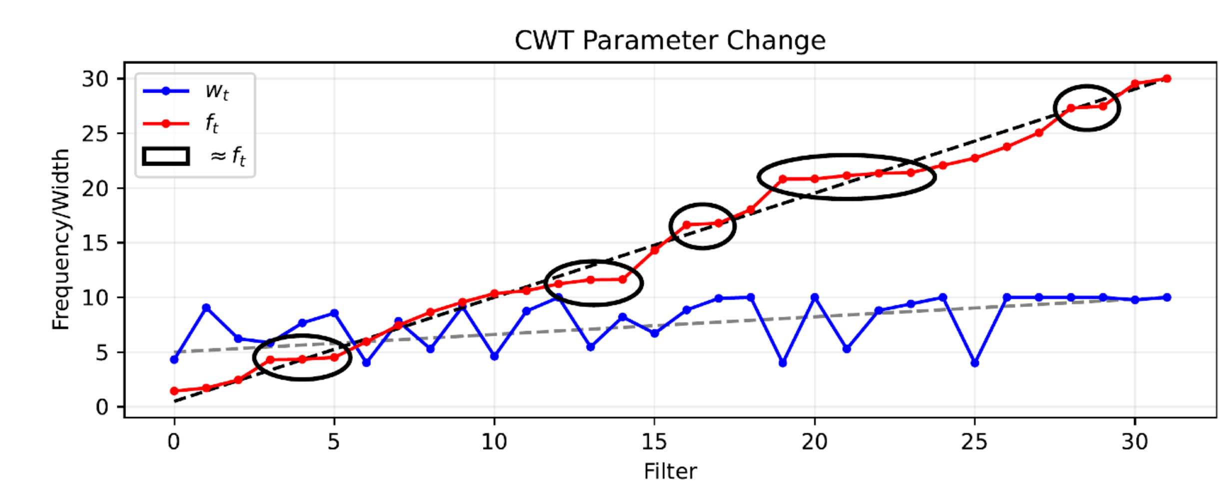

Figure 6: Sorted initial and final frequencies/widths after training.

Discussion

The proposed cwt-net achieves $81.25\,\%$ classification accuracy on the P300 EEG data and is **interpretable by design** with trainable Morlet wavelets that can be be used for time-frequency analysis.

This approach can be extended to other problem domains where a continuous wavelet transform is appropriate and desired for interpretability.

Very **few trainable parameters** are required as with a parameterized function we effectively reduce the number of parameters to learn in the convolutional filter to two (i.e. frequency f and width w).

Future work should consider different architectures following the initial CWT layer and focus on additional interpretability to understand how subsequent layers utilize the magnitude scalogram.