

# Decoding Brainwaves for Visual Attention to Faces and Scenes

Taylor Berger

University of Tennessee, Knoxville  
tberger3@vols.utk.edu

Yiyao Chen

The Chinese University of Hong Kong  
charlotttechen2007@gmail.com

August 3, 2018

## Abstract

Attention is a multidimensional cognitive function that can be broken down into a series of filtering and searching subprocesses. Deficiencies in attention are commonly seen in many brain disorders, such as Alzheimer's Disease and Attention Deficit Disorder (ADD), have an impact on an individual's overall cognitive and perceptual capabilities. Many studies have been conducted to evaluate and train brain computer interfaces (BCI) to distinguish between brain signals released during periods of attention in an individual with average and deficient cognitive ability. This study focused on distinguishing between electroencephalography (EEG) signals released during periods of visual attention to faces and scenes for individuals of normal cognitive ability. Within this work, signal pre-processing methods, such as band-pass filters and the Continuous Wavelet Transform (CWT), and deep learning paradigms were tested and optimized to increase decoding accuracy. Data augmentation was also utilized and tested to increase the model training size to increase decoding accuracy.

**Keywords:** *Brain Computer Interface (BCI), Electroencephalography (EEG), Continuous Wavelet Transform (CWT), Deep Learning, Data Augmentation*

## I. Introduction

In this report, we will outline the purposes and objective of our study, the experiment's data collection process, the signal processing methods, and the deep learning methods used to optimize the model. This paper will also present the results and conclusions about the models developed during this project.

### i. Objective

The goal of the project was to develop a filtering method deep learning model to complete a binary classification of EEG signals collected during a period of visual attention. High testing

accuracy in a deep learning model is critical to the BCI platform's applicability in real-world applications. As a means of optimizing this accuracy, the project objectives are the following:

- Compare signal pre-processing techniques to determine which approach provides the optimal deep neural network input.
- Apply deep learning neural network parameters to optimize the classification testing accuracy.
- Determine the impact of utilizing data augmentation methods to increase the training data size.

## **ii. Background Knowledge: EEG**

Electroencephalogram (EEG) is a monitoring method that collects the spontaneous electrical activity on the scalp of the brain in real time [1]. EEG data is collected through electrodes as a summation of synchronous activity of surrounding neurons with "similar spatial orientation" [2]. Unlike comparative brain monitoring techniques, EEG is a low-cost, non-invasive method, making it an area of intrigue in the development of BCI applications.

## **II. Experimental Methods and Materials**

This section outlines the material and components used in the construct of this BCI platform and the experimental protocol used to collect the EEG data from the participants during the study. The layout of the experimental blocks and image categories are also explained in detail in this section.

### **i. Data acquisition**

The BCI platform for this experiment consisted of a wireless EEG headset, dual PC monitors, and data analysis software. Raw EEG data was collected noninvasively through the 14 channel Emotiv EPOC EEG recording headset. Before each application, channel electrodes were hydrated and examined to ensure quality conductance. The placement of each channel was based on the 10-20 international system. This location system ensures that data from frontal, temporal, and occipital lobe regions of the scalp is collected. The labeled locations of these electrode channels are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. The sampling frequency for the headset was set to 128 Hz for the duration of the experiment. During the data collection process, an initial filtering technique was applied to the raw signal to reduce noise. Each signal was filtered using a high-pass filter of 0.2 Hz and a low-pass filter of 43 Hz. The headset transmits the signals to the collection software through a Bluetooth connection. The data collection software used in this experiment was MATLAB and Simulink. Using this platform, the raw EEG signals were collected and stored as time-series data.

### **ii. Experimental protocol**

The experimental procedures and data acquisition methods for this study were approved by the Institutional Review Board at the University of Tennessee. All 38 study participants self-reported good health and were a mixture of male and female subjects. Each subject voluntarily completed the experiment's eight training blocks. All participants had normal or corrected-to-normal vision and had no history or neurological or psychological disorder (based on self-

report). Before beginning the experiment, written consent was obtained for all participants. To complete each experimental block, a PC with dual monitors was set up in front of the subject and the administrator. One monitor was for the administrator to control the experiment. The second screen was positioned in front of the subject to present the image stimuli. Seated approximately 50 cm in front of the monitor, the subjects positioned themselves with one hand resting on their lap and the other on the computer mouse. The computer mouse was used for indicating behavior responses as a secondary means of ensuring visual attention. Aside from the mouse signals, subjects were instructed to limit excessive body movement, and pay attention to the monitor for the entirety of each phase.

### iii. Experimental tasks

The experiment consisted of eight trial blocks of 50 images per trial. Between each block, subjects received a break to limit the fatigue. Before each block, an instructional slide for image category and subcategories was shown. The 50 image stimuli within each block were an assorted mixture of two subcategories. This experiment used four subcategories: indoor scenes vs. outdoor scenes and male faces vs. female faces. Images used in this experiment were black and white and equal in size. Each image was placed on the screen in front of the subject for a duration of 1000 ms before transitioning into a black blank screen. A 1-1.5 seconds blank screen between each image stimuli allowed the subject to ready themselves for the next image. To prime the subject to the category, participants distinguished between task-relevant images (ex. male) and task-irrelevant images (ex. female). Table I outlines each block's task-relevant and irrelevant subcategories. For task-relevant images, subjects responded with keyboard signal. This experimental model ensured the participant gave visual attention to the image category. Binary classification between image categories was established for each image category based on the hypothesis that brain signals contained common features between image subcategories.

**Table I.** The task-relevant and task-irrelevant images for each experimental block.

Block Number	Task-Relevant Image	Task-Irrelevant Image
1	Indoor	Outdoor
2	Male	Female
3	Indoor	Outdoor
4	Female	Male
5	Outdoor	Indoor
6	Male	Female
7	Outdoor	Indoor
8	Female	Male

### iv. Our Contribution

We have applied signal processing methods and deep learning techniques, such as neural networks, to create a general binary classification model for signal analysis.

## III. Pre-Processing Methodology

This section outlines the multiple signal processing methods and feature extraction techniques tested on the EEG signals. Both the band-pass filter and the Continuous Wavelet Transform (CWT) are used within this study in multiple approaches and are explained in detail in the following section. Subsequently, this section also outlines the use of data augmentation to increase the data input into the model and discusses the biological implications of this method.

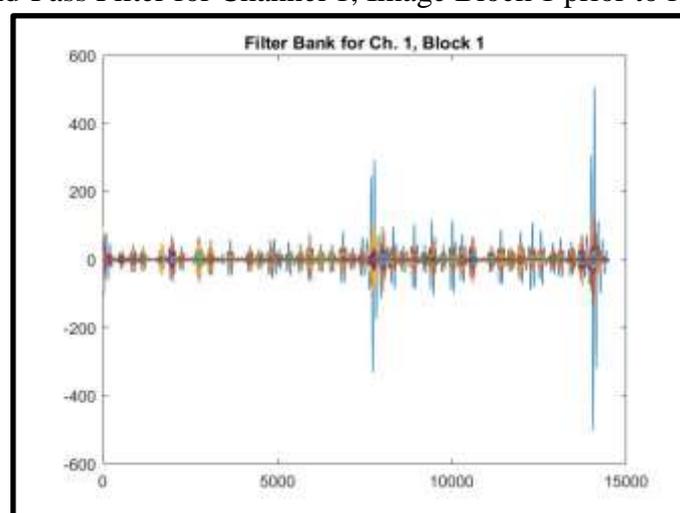
### **i. Pre-Processing**

Due to the nature of EEG signal collection, the recorded signals are prone to various artifacts such as eye blinks and facial twitches. These muscle artifacts create noise in the signal that could hinder the BCI platform's ability to evaluate and train the model. To limit the effects of these movements within the results of the experiment, the signal was analyzed and pre-processed. Within this study, the pre-processing techniques focused on the frequency domain of the signal. The study relied on the five accepted frequency ranges commonly analyzed within the brain: Delta [0.5-3 Hz], Theta [3-8 Hz], Alpha [8-12 Hz], Beta [12-30 Hz], and Gamma [ $>30$  Hz] [3]. A description of the signal processing methods used in this study is given below.

#### **i.1 Band-Pass FIR Filter**

A band-pass FIR filter with a minimum order was run on each channel of the EEG signal [4]. The band-pass filter was set with a high-pass filter of 3 Hz and a low-pass filter of 59 Hz. From this filter, 56 frequency bands were constructed with a width of 1 Hz. For each frequency band, five distinguishing statistics were extracted: mean, kurtosis, variance, skewness, maximum [5]. The output of this filter was fed into a Convolutional Neural Network (CNN) model with an input shape of  $56 \times 14 \times 5$ , where the five distinguishing signal statistics were stacked. Picture I is a visual of the filter output for channel 1, image block 1.

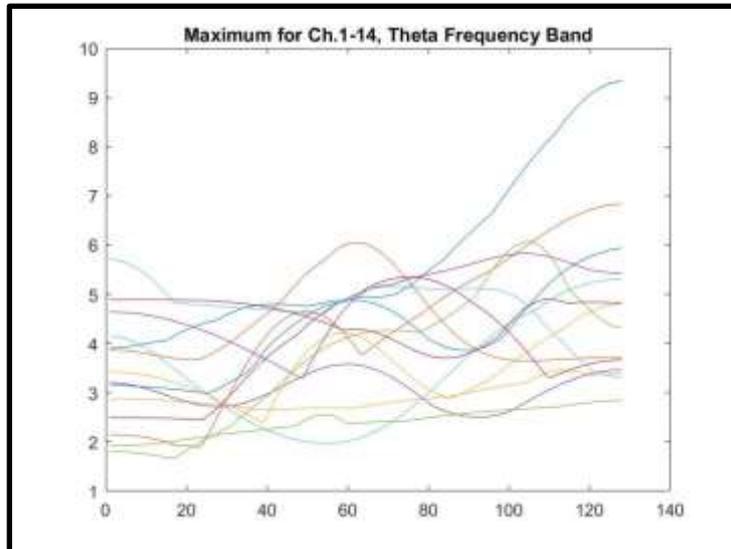
**Picture 1.** Band-Pass Filter for Channel 1, Image Block 1 prior to feature extraction



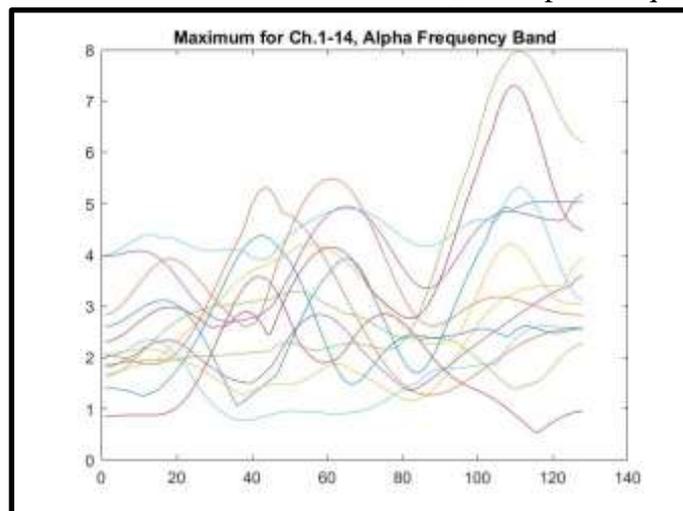
#### **i.2 CWT - Statistical Parameters**

Computed by the analytic Morlet wavelet, the CWT was applied to each channel of the EEG signal [6]. Frequency bands for the Theta, Alpha, Beta, and Gamma frequency ranges were extracted. Within each frequency range, three distinguishing statistics were extracted: mean, maximum, and variance. The output of this filter was input into the CNN model with an input shape of 3x14x4, where the four frequency ranges were stacked in the third dimension. Pictures 2-5 are a visual for the maximums for channels one through 14 in each frequency band.

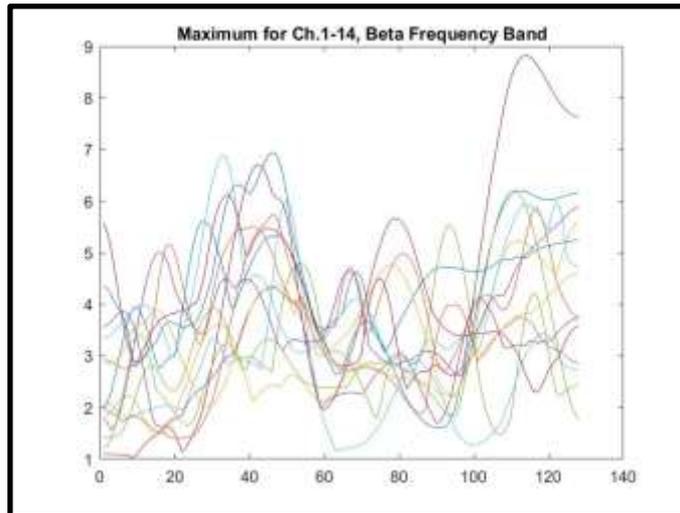
**Picture 2.** Statistical Parameter: Maximum, for Theta Frequency Band



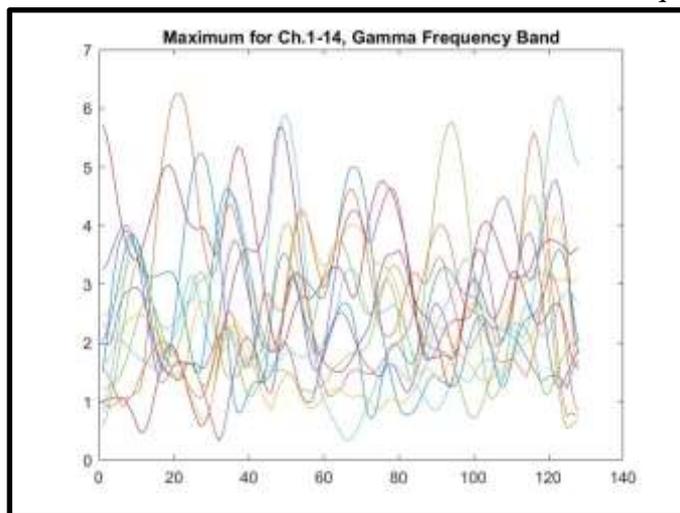
**Picture 3.** Statistical Parameter: Maximum, for Alpha Frequency Band



**Picture 4.** Statistical Parameter: Maximum, for Beta Frequency Band



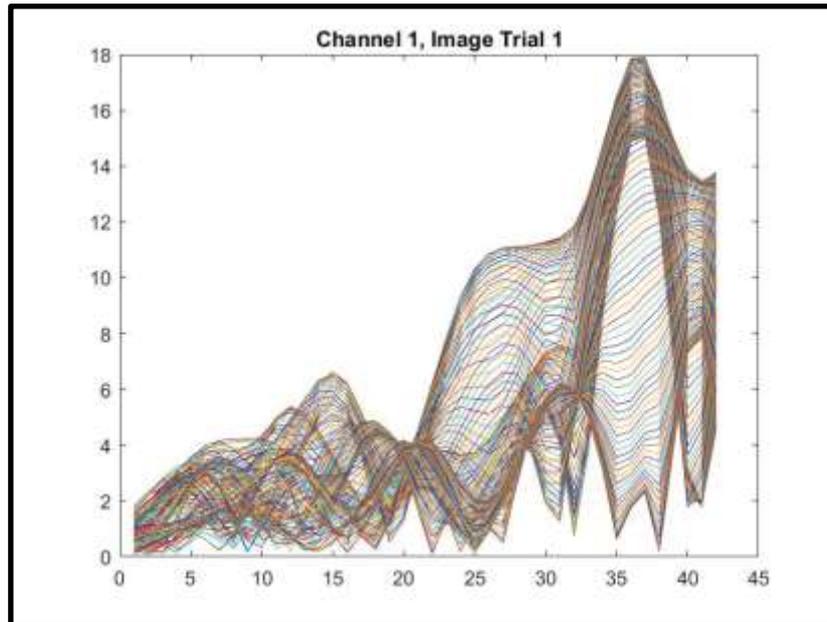
**Picture 5.** Statistical Parameter: Maximum, for Gamma Frequency Band



### **i.3 CWT - 2D Image Input**

The CWT was applied to each channel of the EEG signal by computing the analytic Morlet wavelet. For each channel, 42 signals were decomposed and unaltered prior to being input into the CNN model. The output of this filter was input into the CNN model with an input shape of 42x128x14, where the 14 channels were stacked as images. Picture 6 is a visual of the output array for channel 1, image trial 1.

**Picture 6.** CWT Filter for Channel 1, Image Trial 1



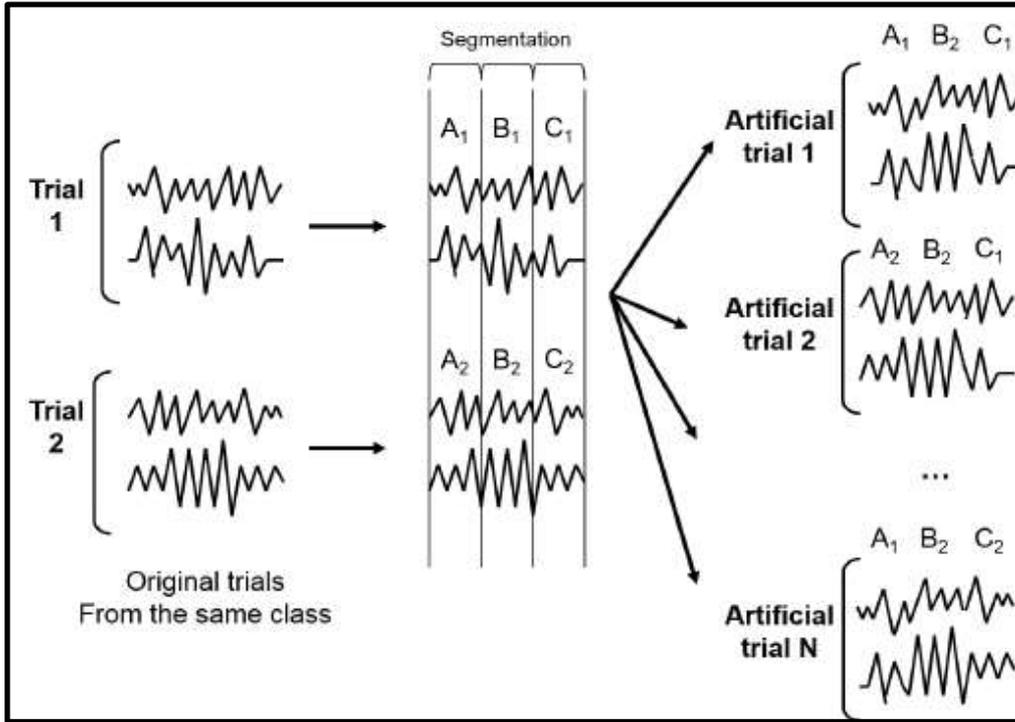
## ii. Data Augmentation

When compared to other fields, BCI platforms have a limited data set on which the model is expected to train and test. These limited training datasets do not utilize the full capabilities of a deep learning model, when compared to models who use huge training sets in order to extract distinguishable features. Following methods that have been previously explored and tested, we have integrated a data augmentation method to increase the available training set for the deep learning model [7]. The split-and-combination method, developed by Fabien Lotte, was applied to EEG signals to generate augmented data [8]. The validity in EEG data augmentation is disputable between neurobiologists and computational sciences. By nature, EEG signals are a considered continuous. As result, each data point within the EEG signal is dependent on the surrounding data points. This method of data augmentation does not preserve the continuous nature of the signal.

### ii.1 Split-and-Combination Method

Data augmentation is applied to the model prior training data set of the model. The training set for each model consists of 360 randomly selected image trials. These 360 image trials are take and divided into the binary image categories, faces (label 1) and scenes (label 0). After this separation, all channel signals within the EEG image trial are cut into  $k$  segments. The number of segments,  $k$ , must be an integer and a factor of the number of sampling points within each signal. In our model, the length of each signal was 128 or 126, dependent on the signal processing method applied. Once the signal was segmented, signal segments were randomly selected from each signal within the image category and combined with  $k$  number of segments in the corresponding order, outline in Picture 7. By applying this method, 600 artificial signals with label 0 and 600 with label 1 are created and therefore 1600 examples are used for model construction.

**Picture 7.** Split-and-Combination Method for EEG Data Augmentation [8]



#### IV. Deep Learning Model

To conduct the binary classification of the collected EEG signals, supervised deep learning models were employed. To determine the best neural network architecture, a Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) were created and implemented using Keras framework. Each model was evaluated with input data from different preprocessing methods previously outlined in the report. For each subject, there are 400 examples in total valid: 200 with label 0 (scene image) and 200 with label 1 (human face image). Models are trained on 360 image trial examples and tested on the remaining 40 image trial examples.

##### i. Raw Signal Modeling

As a means of determining which model architecture to use, a baseline test on the minimally filtered EEG data was conducted. Low-pass filtering with a Nyquist cutoff of 40 Hz was applied to the signal prior to the data being input into the models. The structures of both the CNN and RNN models are described below in detail.

##### i.1 Convolution Neural Network

After running the signal through a low-pass filter, the input shape for each image trial was 128 time points by 14 channels. In order to create a perfect square for the input shape, it was determined that the first two time points would be removed, reducing the image trial shape down to 126 time points by 14 channels. Each image trial could then be reshaped to an input size of 42 by 42. The CNN model outlined in Table II includes two-dimensional convolutional layers with a kernel size of three-by-three. Max-Pooling layers are also incorporated into the

model with strides sizes of two-by-two. Prior to fitting and compiling the model two fully connected layers, with 128 and 2 nodes sequentially, are incorporated. The structure of this model is outlined below:

**Table II.** Baseline CNN Model Structure

Input (42×42)	conv3×3-6	conv3×3-12	conv3×3-24	FC-128	FC-2
	ReLU	ReLU	ReLU	ReLU	Softmax
	Max_Pooling	Max_Pooling	Max_Pooling		

### **i.2 Recurrent Neural Network**

Within this study, the EEG signals were collected as time-series. While CNN models have been a primary approach applied to EEG signals in BCI platforms, many previous studies have employed RNN models as a point of comparison due to their sequential behavior [9]. A simple many-to-one RNN was constructed with an LSTM layer. This model layout was chosen due to its distinguishment as a comparatively more suitable architecture for a longer time-dependent input. After being passed through the low-pass filter, the EEG signal was formatted as 128 time points by 14 channels for each image trial. The RNN model structure is described below:

**Table III.** RNN model structure

<code>model.add(LSTM(128, input_shape=(timesteps, 14)))</code>
<code>model.add(Dense(1,activation='sigmoid'))</code>
<code>model.compile(loss='binary_crossentropy', optimizer='adam',metrics=['accuracy'])</code>

### **i.3 Conclusion**

After running the EEG signals through each of the models, the CNN model gave a classification accuracy of 50%, chance level, and the RNN model gives a classification accuracy of 57.5%. While the RNN model provided the better results on the minimally filtered EEG data, it was decided that the CNN model would be utilized in future model development. Analyzing the results of these models, it was concluded that pre-processing and data augmentation will be important to the optimization of testing accuracy.

### **ii. Cross-Validation & Random Seed**

To ensure reproducible results from the model, random seeding was fixed into the development of the neural network. K-Fold cross validation employs a method of data splitting to ensure that the training data was randomly chosen for each fold to eliminate the possibility of selecting extreme testing data. To make the model more robust, 10-fold cross-validation was applied to evaluate the performance of the model. The training-to-testing ratio was set to 0.9 for each fold.

Each model was evaluated using the average testing accuracy from each fold and standard deviation.

### iii Signal Processing: Band-Pass

This section outlines the CNN modeling development and optimization process for the data output by the band-pass filter. The results of this model are also explained in detail below.

#### iii.1 Input Data & Model Type

As previously outlined in the *Pre-Processing Methodology* section, the data size after filtering the signal through the band-pass filter is a  $56 \times 14 \times 5$  array for each image trial. To prepare the signal for the model, each image trial was reshaped into  $28 \times 28 \times 5$  array. Within this model, data augmentation was applied across all 400 image trials prior to the pre-processing. In total 600 artificial image trials were generated for each face and scene label for a complete image trial count of 1600. The CNN model uses three convolutional layers with a max-pooling size of two-by-two and is followed by two fully connected layers with 56 nodes and 1 node sequentially. Two-dimensional and three-dimensional convolutional layers were tested and evaluated. The number of layers was optimized by testing and comparing the accuracy of models with three to seven layers and different optimizers are tested. The model with the best testing accuracy was a three-dimensional CNN with an Adam optimizer. The structure of this model is outlined below:

**Table IV.** CNN Model Structure for Multiple Extracted Features Input

Input ( $28 \times 28 \times 5 \times 1$ )	conv $3 \times 3 \times 3$ -6	conv $2 \times 2 \times 2$ -12	conv $3 \times 3 \times 3$ -24	FC-56	FC-1
	ReLU	ReLU	ReLU	ReLU	Sigmoid
		Max_Pooling $2 \times 2 \times 2$	Max_Pooling $2 \times 2 \times 2$		

#### iii.2 Results

Different numbers of epochs and models with early stopping were tested and evaluated. After the completion of the tests, it was concluded that the model with 55 epochs and no early stopping was the best structure (see Table V).

**Table V.** Ten-Fold Cross-Validation Prediction Results over Subject 1 for Multiple Extracted Features Input

3D CNN with Early Stopping	Mean (std) = 74.38% (3.61%)
Batch Size = 1    Patience = 10	Max = 80.62%    Min = 66.88%
3D CNN	Mean (std) = 76.19% (6.64%)
Batch Size = 1    Epoch Number = 55	Max = 85.62%    Min = 64.38%

### iii.3 Conclusion

It was determined that the method of data augmentation applied was susceptible to information leaking. As a result, the test accuracy from this model and filtering technique cannot be considered as accurate. To prevent information from leaking, data augmentation in future models should not be integrated across all image trials. For future models, data will be augmented within each fold across the training datasets only. Applying data augmentation in this method will prevent information leaking and should produce more reliable results.

### iv Signal Processing: CWT - Statistical Parameters

This section outlines the CNN model developed for the CWT filter with extracted statistical parameters outlined in the *Pre-Processing Methodology* section of the report. Below is a description of the model development and optimization process. The results of this model are also described below.

#### iv.1 Statistical Parameters Input

After being filtered, the output data is in the shape of a  $3 \times 4 \times 14$  array for each image trial. To prepare the data for CNN model, the first and the 14th channels were removed to create a  $3 \times 4 \times 12$  array for each image trial. When considering which channels to remove, the vulnerability to noise was evaluated. The positioning of the 1st and 14th electrodes left the channels susceptible to eye movement artifacts more so than any other channel. By cutting these two channels out of the model, we reduced the possibility of this type of noise affecting the results of the model. After the two channels were removed, the input array was reshaped to a  $6 \times 6 \times 4 \times 1$  for each image trial. The CNN structure for this data input is depicted in Table VI.

**Table VI.** CNN Model Structure for Three Extracted Features based on CWT Input

Input ( $6 \times 6 \times 4 \times 1$ )	conv $2 \times 2 \times 2$ -6	conv $2 \times 2 \times 2$ -12	conv $2 \times 2 \times 2$ -24	FC-10	FC-1
	Linear	Linear	Linear	Linear	Sigmoid
		Max_Pooling $2 \times 2 \times 2$	Max_Pooling $2 \times 2 \times 2$		

#### iv.2 Results

During the model optimization process, different numbers of epochs and batch sizes were tested with only original data. After the completion of the test, it was concluded that the optimal batch size and epoch number combination for this input was 50 and 100 sequentially as depicted in Table VII.

**Table VII.** Ten-Fold Cross-Validation Prediction Results over Subject 1 for Three Extracted Features based on CWT Input

Batch size =50	Epoch Number = 100		
Mean (std) = 62% (7.81%)	Max = 75%	Min = 47.5%	

**v Signal Processing: CWT - 2D Image**

This section outlines the CNN model developed for the CWT filter for the two-dimensional that was described in detail in the *Pre-Processing Methodology* section of the report. Below is a description of the model development and optimization process. The results of this model are also described below.

**v.1 2D Image Input**

After being filtered, the signal output is in the shape of a 42x128x14 array for each image trial. To prepare the signal data for the model, the first and the last data points were removed from the signal to create an array size of 42x126x14 for each image trial. After these two data points were removed, the array was reshaped to an input size of 42x42x42 for each image trial. These arrays were treated as images and input into three-dimensional CNN model. Different optimizers were tested and applied and the model with the best performing model with an Adam optimizer is described below:

**Table VIII.** CNN Model Structure for CWT Images Input

Input 42x42x42x1	conv 3x3x3 -6	conv 3x3x3 -12	conv 3x3x3 -24	conv 3x3x3 -24	conv 3x3x3 -48	FC-10	FC-1
	Linear	Linear	Linear	Linear	Linear	Linear	Sigmoid
			Max_Pooling 2x2x2		Max_Pooling 2x2x2		

**v. 2 Results**

The model was tested and optimized for different batch sizes and epoch numbers with only original data from subject 1. After the completion of the tests, it was concluded that the optimal model was determined to be a batch size of 180 and 30 epochs. The results are outlined below:

**Table IX.** Ten-Fold Cross-Validation Prediction Results over Subject 1 for CWT Images Input

Batch Size = 180	Epoch Number = 30		
Mean (std) = 71% (6.04%)	Max = 82.5%	Min = 60%	

**V. Conclusion**

Within this section, the pre-processing and deep learning models will be evaluated and compared. The best performing pre-processing method and model will be applied across all 38 subjects. By comparing the mean and standard deviation of each model, the best performing model was the CNN model with the 2D image input. Run on the data output by subject one, the model performed with an average accuracy of 71% and a range of 60% to 82.5%, as depicted in Table IX. This individual model was taken and applied to the remaining 38 subjects as a general model. The results of this test can be seen in Appendix I: Table X.

The model was also evaluated after applying data augmenting to the training dataset within each fold to keep the datasets independent. Due to the new nature of the data, the number of epochs was reduced to 23 to prevent overfitting. Each CWT decomposed wave is split into  $k$  segments ( $k$  is a factor of 126) and randomly pick each segment and combine them together in order. Each label is augmented to produce 600 new image trial arrays. Therefore the model is trained with 1560 examples and tested on 40 examples.

Several factors (1, 2, 3, 6, 7, 9, 14, 18, 21, 42) of 126 were tested in data augmenting method on subject 33 to determine the optimal number of splits. After this test, it was concluded that the optimal number of segment splits was three (see Appendix II: Table XI). After the application of the data augmenting method, the average accuracy over 38 subjects improved by 0.76%. Within this margin of improvement, 24 subjects increased in testing accuracy, two subjects produced the same accuracy, and 12 subjects declined in accuracy. Within this new model, subject 20 saw the most improvement of 4.5% (see Appendix III: Table XII). These results suggest the importance of developing optimized individual models to test the impact of artificial data on each subject.

Different parameters, such as augmenting size, were also evaluated. As augmenting size increased to 1800 segments per image label, the average testing accuracy across all 38 subjects also noticeably increased. The average testing accuracy across all subjects improved by 1.11%, and subject 4 produced the highest individual improvement of 5.25%. In total, 29 subjects produced improvements in their testing accuracy, 1 subject produced the same result, and 8 subjects declined in accuracy. A breakdown of these results is outlined in Appendix IV: Table XIII.

## **VI. Future Work**

Building off the results of this study, future work will be done in the development of the CWT pre-processing technique and the implementation of augmented data. For the new CWT approach, a three-dimensional image will be constructed from the decomposed signal, representative of the RGB matrix, for each channel. This matrix will be treated as an image and fed through the CNN model. Another approach that will be tested is the increased amount of augmented data. As our results depict, increasing the number of augmented signals developed, increased the accuracy of the CNN model. We would like to run further tests to optimize this approach.

## **VII. Acknowledgments**

This project was mentored by Dr. Xiaopeng Zhao, UTK, Soheil Borhani, UTK, and Dr. Kwai Wong, JICS. This project was sponsored by the National Science Foundation through Research Experience for Undergraduates (REU) award, with additional support from the Joint Institute of Computational Sciences at University of Tennessee Knoxville. This project used allocations from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation.

## VIII. Appendix

**Appendix I: Table X.** Ten-Fold Cross-Validation Prediction Results over 38 subjects  
(30 epochs, batch size = 180, Max/Min accuracy in bold )

Subject	Mean (std)	Min - Max	Subject	Mean (std)	Min - Max
1	71% (6.04%)	60%-82.5%	20	67.75% (6.37%)	55.0%-77.5%
2	71.75% (4.88%)	62.5%-80%	21	67.75% (7.45%)	57.5%-85.0%
3	62.5% (5.92%)	55%-70%	22	68% (6%)	57.5%-77.5%
4	70.25% (5.64%)	65%-77.5%	23	69.5% (5.57%)	60%-77.5%
5	76.5% (5.15%)	65.0%-82.5%	24	64% (6.34%)	57.5%-80%
6	71% (4.36%)	65%-77.5%	25	69% (5.5%)	62.5%-82.5%
7	59.75% (7.02%)	50%-70%	26	72.50% (6.22%)	62.5%-80.0%
8	66.25% (8.46%)	52.5%-77.5%	27	76.75% (5.92%)	67.5%-85%
9	71% (7.09%)	57.5%-80%	28	67.25% (8.25%)	50.0%-80.0%
10	63.25% (7.99%)	55%-75%	29	67.75% (5.41%)	60%-75%
11	66.75% (4.88%)	57.5%-75%	30	75.5% (5.45%)	67.5%-85%
12	64.25% (5.37%)	57.5%-72.5%	31	60.25% (6.17%)	55%-72.5%
13	74.5% (5.45%)	65%-85%	32	69% (4.5%)	62.5%-77.5%
14	72.25% (8.25%)	57.5%-85%	<b>33</b>	<b>79.5% (6.3%)</b>	67.5%-87.5%
15	64.5% (5.89%)	57.5%-75%	34	63% (6%)	57.5%-72.5%
<b>16</b>	<b>59.0% (8.46%)</b>	45.0%-75.0%	35	66.75% (8.44%)	55%-80%
17	66.0% (5.15%)	60.0%-75.0%	36	62% (7.73%)	52.5%-72.5%
18	63.75% (6.54%)	55.0%-77.5%	37	62.5% (9.08%)	47.5%-77.5%
19	63.5% (9.63%)	45.0%-80.0%	38	67.75% (5.96%)	60%-80%
Mean: 67.74%		Max: 79.5%		Min: 59%	

**Appendix II: Table XI.** Ten-Fold Cross-Validation Prediction Results for **subject 33** with Artificial Data

( 23 Epochs, Batch Size = 400, Augmenting Size = 600, Max improvement in bold)

#Segment	Mean (std)	#Segment	Mean (std)
1	79.50% (6.30%)	9	78.25% (7.08%)
2	79.50% (7.40%)	14	77.75% (6.75%)
3	<b>81.25% (5.15%)</b>	18	78.25% (7.59%)
6	77.75% (5.86%)	21	79.25% (6.03%)
7	78.25% (6.43%)	42	77.75% (5.96%)

**Appendix III: Table XII.** Ten-Fold Cross-Validation Prediction Results over 38 subjects with Artificial Data

( 23 Epochs, Batch Size = 400, Segment number =3, Augmenting Size = 600, Subjects with improvement in bold)

Subject	Mean (std)	Subject	Mean (std)
1	69.5%(5.1%)	20	<b>72.25%(5.75%)</b>
2	<b>72.75%(4.8%)</b>	21	<b>67.75%(3.61%)</b>
3	61.75%(4.34%)	22	66.5%(3.74%)
4	<b>73%(6.2%)</b>	23	69%(6.73%)
5	76.25%(5.62%)	24	62.75%(6.66%)
6	<b>73.25%(4.48%)</b>	25	<b>72%(4.85%)</b>
7	58.5%(6.73%)	26	<b>74.75%(6.27%)</b>
8	66%(7.92%)	27	76%(5.39%)
9	<b>72%(4.44%)</b>	28	<b>67.5%(10.84%)</b>
10	<b>66%(6.14%)</b>	29	<b>68%(7.4%)</b>
11	<b>68%(6.5%)</b>	30	<b>77.5%(6.52%)</b>
12	61%(4.9%)	31	60%(5%)
13	<b>75.5%(3.67%)</b>	32	<b>70%(6.61%)</b>
14	<b>72.25%(6.37%)</b>	33	<b>81.25%(5.15%)</b>
15	<b>64.75%(4.93%)</b>	34	<b>63.75%(5.84%)</b>
16	<b>60%(10.37%)</b>	35	<b>68.25%(5.92%)</b>
17	<b>66.5%(5.39%)</b>	36	<b>64.75%(5.53%)</b>
18	<b>65.25%(4.53%)</b>	37	<b>65%(7.83%)</b>
19	66.25%(8.08%)	38	67.5%(6.12%)
Mean: 68.5%		Min-Max: 58.5%-81.25%	

**Appendix IV: Table XIII.** Ten-Fold Cross-Validation Prediction Results over 38 subjects with Artificial Data

( 10 Epochs, Batch Size = 400, Segment number =3, **Augmenting Size= 1800**, Subjects with improvement in bold)

Subject	Mean (std)	Subject	Mean (std)
1	67.25% (2.84%)	20	<b>71.75% (6.71%)</b>
2	<b>73.75%(4.64%)</b>	21	66.75% (5.48%)
3	<b>63.25%(5.60%)</b>	22	<b>69.25% (6.62%)</b>
4	<b>75.50% (5.34%)</b>	23	68.50% (8.67%)
5	<b>78.50% (6.91%)</b>	24	62.25% (7.54%)
6	<b>73.25% (6.62%)</b>	25	<b>70.50% (6.60%)</b>
7	<b>61.50% (8.23%)</b>	26	<b>74.25% (5.25%)</b>
8	<b>68.75% (9.10%)</b>	27	<b>77.00% (4.30%)</b>
9	71.00% (5.15%)	28	<b>69.50% (7.73%)</b>
10	<b>67.00% (6.10%)</b>	29	<b>68.25% (7.16%)</b>
11	<b>68.50% (4.50%)</b>	30	<b>77.50% (5.00%)</b>
12	<b>65.25% (5.96%)</b>	31	<b>61.00% (5.27%)</b>
13	73.75% (3.91%)	32	66.00% (4.21%)
14	<b>72.75% (6.27%)</b>	33	<b>80.00% (5.70%)</b>
15	63.75% (4.91%)	34	<b>65.75% (7.08%)</b>
16	<b>61.50% (9.03%)</b>	35	<b>67.75% (6.93%)</b>
17	<b>66.75% (3.17%)</b>	36	<b>62.75% (5.64%)</b>
18	<b>67.75% (5.18%)</b>	37	62.00% (6.96%)
19	<b>67.50% (8.29%)</b>	38	<b>68.25% (6.90%)</b>
Mean: 68.85%		Min-Max: 61% - 80%	

## IX. References

1. Malmivuo, J., & Plonsey, R. (1995). *Bioelectromagnetism: Principles and applications of bioelectric and biomagnetic fields*. New York: Oxford Univ. Press.
2. Desai, J. (2014). *Electroencephalography(EEG) Data Collection and Processing through Machine Learning*. University of Arkansas, Fayetteville
3. Buzsáki, György (2006). *Rhythms of the Brain*. New York: Oxford University Press. p. 114
4. S. S. Daud and R. Sudirman, "Butterworth Bandpass and Stationary Wavelet Transform Filter Comparison for Electroencephalography Signal," *2015 6th International Conference on Intelligent Systems, Modelling and Simulation*, Kuala Lumpur, 2015, pp. 123-126. doi: 10.1109/ISMS.2015.29
5. Kalpana, R., & Chitra, M. (2012). Classification of Different Brain States From EEG Signals Using Statistical Parameters. *International Journal of Advanced Computing, Engineering, and Application*, 1(3). Retrieved from <http://www.ircast.org/ijacea/papers/vol1no32012/6vol1no3.pdf>
6. Prince, P., & Rani Hemamalini, R. (2012). Seizure detection using parameter estimation and Morlet wavelet transform. *Communications in Computer and Information Science*, 270(II), 674-679.
7. Zhang, Q., & Liu, Y. (2018). Improving brain computer interface performance by data augmentation with conditional Deep Convolutional Generative Adversarial Networks.
8. Lotte, F. (2011). Generating Artificial EEG Signals To Reduce BCI Calibration Time. *5th International Brain-Computer Interface Workshop*, 5th International Brain-Computer Interface Workshop, 2011.
9. Fedjaev, J. (2017). *Decoding EEG Brain Signals using Recurrent Neural Networks*. Technical University of Munich, Germany.