Using Machine Learning to Model EEG-Derived Brain Activity During Emotion Regulation

Mahdis Hojjati (EMBS Student Member, IEEE)*, Shyamal Y. Dharia (EMBS Student Member, IEEE)*, Sergio G. Camorlinga (EMBS Member, IEEE)*, Stephen D. Smith*†, Amy S. Desroches*†, Bronte D. Brenneman†
* *Applied Computer Science*, *University of Winnipeg*, Winnipeg, Canada
† *Psychology*, *University of Winnipeg*, Winnipeg, Canada

***Abstract*—Emotion Regulation (ER) is the ability to manage emotional responses. ER is important for maintaining mental health and handling social interactions, especially under stress. This study explores the brain activity involved in ER using electroencephalography (EEG) and machine learning (ML) models to predict successful and unsuccessful ER. Study participants viewed emotional and neutral images under two conditions: regular viewing and being asked to reduce their emotions. At the end of each experimental trial, participants rated the intensity of their emotional response to the image. Ratings of low intensity (1 and 2) were classified as successful ER, whereas ratings of high intensity (3 and 4) were considered indicative of unsuccessful ER. EEG signals were analyzed in both time and frequency domains to identify patterns linked to ER. In the time domain, significant differences in Global Field Power (GFP) were observed, especially in the frontal and central regions of the brain. Frequency-domain analysis using Power Spectral Density (PSD) showed that theta, beta, and gamma bands were important for regulating emotions. Using these analysis results, machine learning models were trained to predict regulation success. Among the models, a neural network with Maximum Mean Discrepancy (MMD) loss performed the best, achieving an F1-score macro of 75.57% with a subject-independent approach. These machine-learning models highlight the importance of frontal and central brain regions and beta brain frequency signals in the prediction of ER levels. It shows that combining EEG data with advanced machine learning methods can create accurate models for understanding and predicting emotional responses. Additionally, this integrated EEG-based approach represents a novel framework for ER assessment, offering a promising direction for future research and enabling personalized mental health treatments.**

***Index Terms*—Emotion Regulation (ER), Electroencephalography (EEG), Machine Learning, Neural Networks, Global Field Power (GFP), Power Spectral Density (PSD)**

I. Introduction

Emotion Regulation (ER) is the ability to manage and modify emotional responses to external events and experiences as well as to one's own thoughts [1]. ER helps us maintain control over our feelings and actions, thereby increasing the likelihood that we will generate situation-appropriate responses. Previous research has identified numerous ER strategies including the reappraisal of emotional thoughts or situations, shifting or distracting attention away from emotional stimuli, and actively suppressing emotional reactions [1]. Although the effectiveness of these strategies varies, research typically shows that ER is an important contributor to both mental health and effective social interactions [2]. Given this importance, it is imperative for researchers to identify the neural structures that work together to perform ER. Doing so will provide important insights into the cause of ER failures in different populations, such as the elderly or individuals with neurological disorders [3], and may eventually allow us to predict when such emotion dysregulation will occur.

The popular neuroimaging technique, electroencephalography (EEG), is useful for understanding ER neural processes. EEG can be used to assess how ER processes can alter brain responses to visual stimuli, such as emotional images or videos. This experimental setup has been widely used, forming the foundation for EEG-based emotion recognition studies [4]–[6]. EEG technology has proven effective for emotion recognition, particularly when combined with advanced machine learning models [7]–[10]. These models enhance the accuracy of detecting human emotions by analyzing the spectral, spatial, and temporal features of EEG signals. However, to our knowledge, no previous research has used machine learning models to examine the patterns of neural activity associated with ER.

To that end, this study aims to collect and analyze brain signals when participants are specifically asked to reduce their emotional responses. Collected EEG signals are then used to identify neural patterns related to ER. After each stimulus, participants were asked to input the intensity of their emotional response after regulation processes on a scale of 1 (low) to 4 (high); this helped measure subjective emotional intensities along with regulation processes. Finally, machine learning models were developed to predict subjective emotional ratings after participants attempted to reduce their emotional response to a stimulus.

II. Related Works

A. Emotion Regulation using EEG

Numerous studies have used EEG to investigate the neural patterns associated with ER. For instance, Kosonogov et al. [11] examined the effects of different ER strategies on brain

This work was supported by Misericordia Health Center, Mitacs Accelerate (IT39756), Natural Sciences and Engineering Research Council of Canada (RGPIN-2023-03443).
Corresponding authors: s.camorlinga@uwinnipeg.ca, s.smith@uwinnipeg.ca, a.desroches@uwinnipeg.ca

Manuscript 1288 submitted to 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) . Received February 7, 2025.

activity by analyzing EEG-based indices of engagement and arousal. These indices were derived from spectral power in specific frequency bands, providing insights into how cognitive and emotional processes are modulated during ER tasks. Furthermore, Del Popolo Cristaldi et al. [12] used EEG to study how ER strategies affect the brain's handling of emotions. They found that ER did not influence how the brain first predicts emotions. However, expressive suppression activated brain areas linked to self-control, while cognitive reappraisal boosted activity in regions that help adjust emotions based on new information. These results show that different ER strategies shape brain responses in changing situations. Additionally, Thiruchselvam et al. [13] found that distraction reduced the late positive potential (LPP) faster than reappraisal during regulation. Later, when the participants were shown the images again, those that had been previously associated with distraction caused a stronger late positive potential (LPP) compared to images that were simply viewed without any regulation. This suggests that distraction and reappraisal affect different parts of the emotional process. Finally, Dennis et al. [14] found that higher frontal EEG activity during mood inductions (fearful, sad, or neutral) compared to a resting baseline was linked to better ER and participants with greater frontal activity reported less sadness and anxiety after the mood inductions. Similar to [13] our study uses LPP and the ER suppression strategy. Unlike [14] which was focused specifically on the frontal lobe, we aim to identify neural patterns across multiple brain regions when individuals are instructed to reduce their responses to an upcoming visual stimulus. Crucially, our approach integrates both time-domain and frequency-domain analyses to uncover robust neural biomarkers predictive of successful ER.

B. Emotion Recognition using Machine Learning

In recent years, many research has focused on developing advanced machine learning models for emotion recognition using EEG systems [7], [15], [16]. A large portion of this work leverages well-established datasets such as the SEED datasets [4], [5], [7], [17], where participants are exposed to audio-visual stimuli designed to elicit emotional responses. These studies have achieved notable success in terms of classification accuracy, demonstrating the potential of machine learning in decoding emotions from EEG signals.

Current SEED datasets and state-of-the-art models perform well for emotion classification. However, these models do not use emotional intensity in their prediction. Dharia et al [18] proposed a framework for classifying subjective emotional intensity during simple emotional perception. The present study building upon [18], advances the field to further understand the neural mechanisms of emotional regulation by predicting various levels of regulation, rather than focusing purely on emotion classification.

III. Methodology

A. Dataset

The dataset includes EEG recordings from 21 university students who participated in an experiment aimed at examining the regulation of emotional responses to visual stimuli. This involved presenting participants with 168 color photographs that were used in our previous studies [18] and [19] categorized into emotional and neutral groups. During the experiment, for half of the images, participants were asked to actively reduce their emotional responses, while for the other half, they were instructed to simply view the images normally without regulating their responses (normal viewing). The participants subsequently rated the intensity of their emotional response on a scale from 1 (minimal emotional response) to 4 (maximum emotional response). Figure 1 shows the whole process in detail.

EEG data were collected using the OpenBCI Cyton Daisy Board in combination with OpenBCI Gelfree Electrode cap [1]. Additionally, a Photodiode Module was used for precise event marking, ensuring synchronization of the stimuli presentation and EEG recording. Data collection was carried out in a single session divided into three blocks, with rest periods between each to minimize fatigue and maintain focus.

The participants, primarily young adults, ranged in age from 16 to 36 years, with an average age of 22.1 and a standard deviation of 4.9. The gender distribution included 14 females and 7 males. Ethical considerations were rigorously adhered to, with all participants providing informed, written consent before participation. The study's protocols conformed to the ethical standards of the institutional research ethics board and the national research committee, as well as the 1964 Helsinki Declaration and its later amendments, ensuring the protection of participant rights throughout the study.

For data analysis, we focus on negative images that evoked strong emotional responses. To maintain the integrity of the experimental paradigm and mitigate participant fatigue, neutral images were also incorporated as control stimuli within the experiment.

B. Preprocessing Pipeline for EEG Data

The preprocessing pipeline used in this study was adapted from the methodology detailed in [18]. Initially, raw EEG data files were loaded for each participant. We made sure that the sampling rate was set to 125 Hz for all participants. Figure 2 shows all the electrodes placements used in data collection.

Further, the signals were filtered using a high-pass filter at 0.01 Hz to eliminate low-frequency drifts and a low-pass filter at 62 Hz to attenuate high-frequency noise, including muscle artifacts and environmental interferences. A notch filter at 60 Hz was applied to remove power line noise. As previously used in [18], [19] in which authors used a temporal lobe site as reference, our online reference channel during data collection was T3 near to the earlobe, and the ground electrode was placed near to Fz. The data was then segmented into

[1] https://openbci.com

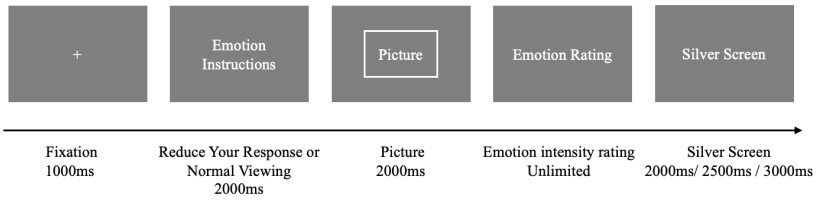


Fig. 1: Initially, participants focus on a fixation point for 1000 ms to ensure attentional readiness. Following this, an instruction is given prior to the stimulus, indicating whether participants should view the upcoming stimulus as "Normal Viewing" or "Reduce Your Response" for 2000 ms. Then a picture stimulus is presented for 2000 ms, during which emotional and cognitive responses are elicited. Immediately after the stimulus presentation, participants rate their emotional intensity on a scale from 1 (low) to 4 (high). Each experimental block consists of 56 trials, with participants completing a total of 168 trials throughout the study.

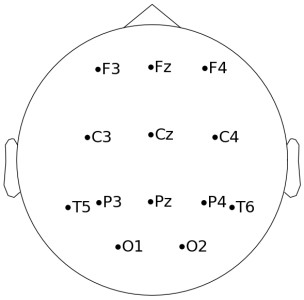


Fig. 2: Electrodes placement used for data collection.

epochs spanning from -400 ms to 2 seconds relative to stimulus onset. This extended duration was chosen to capture delayed emotional regulation effects, which may not appear immediately after stimulus presentation.

An automated artifact rejection procedure, using AutoReject [20], was employed to detect and interpolate bad segments within the epochs. Following this, Independent Component Analysis (ICA) was performed to identify and remove components associated with ocular artifacts. To detect eye blinks, electrooculographic (EOG) epochs were created from F3 and F4, channels located near the eyes. ICA components were evaluated for correlation with these EOG epochs, with a threshold set at 0.5 for the maximum absolute score across the EOG channels. Components exceeding this threshold were classified as ocular artifacts and removed, resulting in ICA-cleaned epochs. After ICA cleaning, a second round of automated artifact rejection was applied. Baseline correction was performed using the pre-stimulus interval from -400 ms to 0 ms, and epochs with amplitudes exceeding ± 100 μV were discarded. Out of the original 3525 epochs, 2424 epochs (68.77%) were retained.

C. Time Domain Analysis

The first stage of the analysis was conducted in the time domain, focusing on identifying differences in the EEG signals over time across the two experimental conditions. This analysis utilized the Global Field Power (GFP) as a measure of the spatial variability of EEG signals over the scalp. GFP provides insights into the dynamics of the brain's global electric field and is particularly sensitive to condition-dependent changes, by capturing the global variance of neural activity across all electrodes [21].

The GFP is mathematically defined as:

$$\text{GFP}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(V_i(t) - \overline{V}(t)\right)^2} \tag{1}$$

Where:

- $V_i(t)$ represents the voltage recorded at electrode i at time t,
- $\overline{V}(t) = \frac{1}{N}\sum_{i=1}^{N} V_i(t)$ is the mean voltage across all electrodes at time t,
- N denotes the total number of electrodes.

Using the conditions that we explained in Section III-A, the overall flow of the analysis is outlined in Algorithm 1:

Algorithm 1 Time Domain Analysis

Input: Two experimental conditions C_1 and C_2, Preprocessed EEG data for all participants
for each participant do
  for each condition C_1 and C_2 do
    # Identify Epoch Responses:
    Extract epochs corresponding to the condition
    if fewer than one epoch is available **then**
      Remove the participant from the analysis
    end if
    # Compute GFP:
    Average epochs to obtain a single evoke signal
    Compute GFP for the evoke signal using Eq. (1)
  end for
end for
 # Perform Statistical Analysis:
Conduct a repeated-measures ANOVA on GFP values of all participants to determine significant differences between C_1 and C_2

D. Frequency Domain Analysis

Following the time-domain analysis, we extended the analysis to the frequency domain. This step aimed to explore differences in brain activity across predefined frequency bands and brain regions. For this analysis, EEG data were grouped into five frequency bands: delta (0-4 Hz), theta (4-8 Hz),

alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-45 Hz). Additionally, data were examined across specific brain regions: frontal (F3, F4, and Fz), central (C4, C3, and Cz), parietal (P3, P4, and Pz), occipital (O1 and O2), and temporal (T5 and T6).

The methodology followed a similar structure to the time-domain analysis mentioned in Algorithm 1, with one key distinction: instead of calculating the GFP, we computed the Power Spectral Density (PSD) for each evoke in step 11. Subsequently, repeated-measures ANOVA tests were performed on the PSD values to determine significant differences.

The PSD was computed as follows:

$$\text{PSD}(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) e^{-j2\pi f n/N} \right|^2 \tag{2}$$

Where:

- N: The total number of samples in the signal $x(n)$,
- $x(n)$: The discrete-time signal being analyzed,
- f: The frequency at which the Power Spectral Density is evaluated,
- $e^{-j2\pi fn/N}$: The complex exponential term representing the Fourier Transform basis function, where j is the imaginary unit.

E. Feature Extraction and Normalization

From the cleaned EEG epochs, we categorized the ER data into two experimental conditions: Reduce Response 1+2 and Reduce Response 3+4. The Reduce 1+2 trials represent cases in which participants successfully reduced their emotional responses to stimuli, whereas Reduce 3+4 trials represent cases where stimuli elicited strong emotions despite participants being told to regulate their responses. For consistency, only participants with sufficient data in both conditions were retained for analysis, totaling 16 subjects.

To minimize noise and variability in the data, epochs were grouped and averaged into sets of five, which we call **super-epochs**. This method offers several advantages. First, it reduces processing time. For example, if we are testing the model on a new participant, instead of running the entire experiment and collecting a large amount of data, we would only need to show five stimuli and use their average as input to the model. Second, this approach is consistent with event-related potentials (ERPs) experiments, where ERPs are obtained by averaging EEG signals for a particular condition. From a practical standpoint, providing a single epoch might introduce errors; however, averaging five epochs results in a more reliable event representation. Finally, this method helps address the data imbalance problem we faced, as grouping and averaging epochs ensures a more consistent representation of the data. This procedure was applied independently for each condition and participant. The PSD for each super-epoch was then computed, providing a concise representation of the signal's frequency-domain characteristics. The resulting PSD features served as the foundation for subsequent machine learning models aimed at classifying the two experimental conditions.

Finally, features were scaled using Min-Max normalization, ensuring the features were scaled in the range of 0 to 1:

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{3}$$

Where X is the original value, $X_{\min}$ and $X_{\max}$ are the minimum and maximum values found in the dataset.

Min-Max scaling was applied to normalize the feature values separately for the training and test sets.

F. Support Vector Machine

We used Support Vector Machine (SVM), which is a traditional ML model to compare its performance with Neural Network (NN) models. For hyper-parameter tuning, a subset of three participants was randomly selected. The data was divided into a training set, consisting of data from two participants, and a validation set, consisting of data from the third participant. The optimal hyper-parameters were determined based on the F1-score macro obtained on the validation set. Once identified, the best parameters were applied to train the SVM model on the remaining subjects' data.

G. Neural Network architecture

Our neural network architecture consisted of four fully connected layers with 128, 64, 32, and 2 units, respectively, each followed by ReLU activation functions and dropout layers with a rate of 0.2 for regularization. The categorical cross-entropy loss was used as the objective function, and the AdamW optimizer was employed with a learning rate of 0.001 and a weight decay of 0.01.

To address class imbalances, a weighted cross-entropy loss was employed, with class weights set to 0.5 for class 0 (Reduce response 1+2) and 1.0 (Reduce response 3+4) for class 1, meaning the weight for class 1 was twice that of class 0. The loss function used is the cross-entropy loss, defined as:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log(\hat{y}_{i,c}) \tag{4}$$

Where, $y_{i,c}$ is the ground truth label, $\hat{y}_{i,c}$ is the predicted probability for class c, N is the number of samples, and C is the number of classes.

H. Neural Network architecture with MMD Loss

Building upon the previous neural network, Maximum Mean Discrepancy (MMD) loss was also used to reduce domain shifts between source (training) and target (test) and to improve generalization across subjects. The final objective function combines the categorical cross-entropy loss and the MMD loss, weighted by a factor α set to 0.1.

The overall loss is defined as:

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \alpha \mathcal{L}_{\text{MMD}} \tag{5}$$

Where:

$$\mathcal{L}_{\text{MMD}} = \left\| \frac{1}{n} \sum_{i=1}^{n} \phi(x_i^{\text{source}}) - \frac{1}{m} \sum_{j=1}^{m} \phi(x_j^{\text{target}}) \right\|^2 \tag{6}$$

Here, ϕ represents the feature mapping function, x_i^{source} and x_j^{target} denote samples from the source and target domains respectively, and n and m are the numbers of samples in each domain.

I. Evaluation of the Models

All models, SVM, Neural Network, and Neural Network with MMD loss, were evaluated using Leave-One-Subject-Out Cross-Validation (LOSO CV), ensuring subject-independent testing. The evaluation metric used was the F1-score macro, averaged across all participants, providing a robust and comprehensive measure of the model's generalization capability.

IV. Experiments and Results

A. Time Domain Analysis

The time-domain analysis results, as shown in Table I, indicate significant differences in the GFP signals when comparing two conditions: viewing negative images while reducing emotional responses versus viewing negative images without regulation. These differences were predominantly observed in the central and temporal lobes, with a p-value < 0.05 using repeated-measures ANOVA. Additionally, since previous work [14] suggests the influence of frontal site for ER we combined the frontal and central regions and we conducted a repeated-measures ANOVA. This analysis also yielded a significant p-value < 0.05, emphasizing the critical role of these brain regions in ER. This combined GFP result is visually represented in Figure 3a. Furthermore, it was observed that brain activity (voltage) was higher when participants attempted to regulate their emotions compared to when they passively viewed the images.

To further investigate the ability to reduce emotional responses to negative images, we compared conditions where participants were instructed to regulate their emotions. After each trial, participants rated the stimuli based on how they felt following the regulation attempt. Ratings of low intensity (1 and 2) indicated successful ER, while ratings of high intensity (3 and 4) suggested less effective regulation. Table II presents the repeated-measures ANOVA results for ER across different brain regions. Significant differences were observed between low-intensity and high-intensity in all brain regions except the parietal lobe, which did not exhibit significant differences. Notably, the frontal and central regions showed the most pronounced effects, highlighting their critical role in ER. Figure 3b shows the GFP comparison between these two conditions using frontal and central electrodes.

Interestingly, the high-intensity condition exhibits higher GFP compared to the low-intensity condition. This observation can be interpreted in two ways. First, the emotional intensity of the image may have been too strong, rendering the regulation process unsuccessful, and leading participants to experience high-intensity emotions. Alternatively, the participants' attempts at ER may have been insufficient, causing them to experience heightened emotional responses despite their efforts.

TABLE I: Time Analysis Using ANOVA (Negative Reduced vs. Negative Normal viewing)

Region	ANOVA Results
Frontal (F3, F4, Fz)	$F(20) = 1.7706, p = 0.1983$
Central (C3, C4, Cz)	$F(20) = 7.1780, p = 0.0144^{*}$
Parietal (P3, P4, Pz)	$F(20) = 0.4491, p = 0.5104$
Occipital (O1, O2)	$F(20) = 1.7136, p = 0.2054$
Temporal (T5, T6)	$F(20) = 6.4947, p = 0.0191^{*}$
Frontal & Central	$F(20) = 5.98, p = 0.0238^{*}$

Note: $^{*}p < 0.05$

TABLE II: Time Analysis Using ANOVA (Negative Reduced 1+2 vs. 3+4)

Region	ANOVA Results
Frontal (F3, F4, Fz)	$F(15) = 22.42, p = 0.0003^{***}$
Central (C3, C4, Cz)	$F(15) = 11.98, p = 0.0035^{**}$
Parietal (P3, P4, Pz)	$F(15) = 3.30, p = 0.0893$
Occipital (O1, O2)	$F(15) = 6.33, p = 0.0238^{*}$
Temporal (T5, T6)	$F(15) = 8.80, p = 0.0096^{**}$

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

B. Frequency Domain Analysis

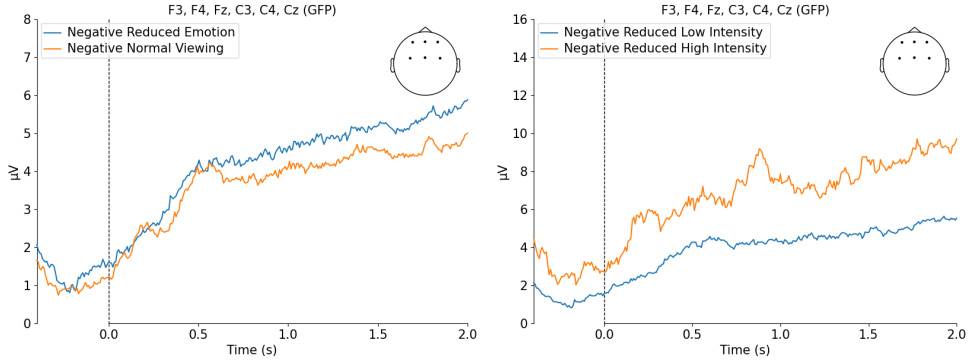
In the frequency-domain analysis, Table III highlights the significant differences between low intensities (1+2) and high intensities (3+4) reduced response in PSD across frequency bands and brain regions, determined by repeated-measures ANOVA test. The results indicate significant differences across all brain regions, specifically in the theta, beta, and gamma frequency bands. Notably, the frontal region exhibits significant differences across all the frequency bands with the beta band showing the highest difference $p < 0.01$. Additionally, Figure 4 illustrates the directional differences in power between these conditions for frontal and central brain regions.

C. Machine Learning

To evaluate the predictive capability for ER, we defined binary classes: **Class 0** - Successful ER (Negative Reduction levels 1 and 2), and **Class 1** - Unsuccessful ER (Negative Reduction levels 3 and 4). The dataset consisted of 127 samples for class 0 and 42 samples for class 1, resulting in an imbalanced class distribution. This imbalance was addressed by using super-epochs, where the average response of 5 epochs was combined into one super-epoch for each subject separately.

Table IV presents the results obtained when training all three models using all EEG channels while selecting only one frequency band as a feature. The beta frequency band consistently achieves the highest F1-score (macro) across all three models and frequency bands. Among the models, the Neural Network consisting of four fully connected layers with MMD Loss performs the best, achieving an F1-score (macro) of **74.58%**, outperforming both the standard Neural Network and SVM. These results support our frequency-domain analysis and confirm the beta band's critical role in enhancing model performance for effectively recognizing successful or unsuccessful ER.

Based on our statistical analysis in the time domain and frequency domain shown in Table II and III, significant



(a) (b)

Fig. 3: Comparative GFP analysis for ER under different conditions. (a) Shows the time-domain analysis for regulated versus unregulated emotional responses to negative images. (b) Displays the GFP comparison between low-intensity and high-intensity ER conditions.

TABLE III: Frequency Analysis Using ANOVA (Negative Reduced 1+2 vs. 3+4)

Region	Delta	Theta	Alpha	Beta	Gamma
All regions	$F(15) = 10.00, p = 0.006^{*}$	$F(15) = 9.18, p = 0.009^{*}$	$F(15) = 3.44, p = 0.083$	$F(15) = 7.18, p = 0.017^{*}$	$F(15) = 9.77, p = 0.007^{*}$
Frontal (F3, F4, Fz)	$F(15) = 10.01, p = 0.006^{*}$	$F(15) = 8.98, p = 0.009^{*}$	$F(15) = 6.21, p = 0.025^{*}$	$F(15) = 11.34, p = 0.004^{**}$	$F(15) = 7.02, p = 0.018^{*}$
Central (C3, C4, Cz)	$F(15) = 5.30, p = 0.036^{*}$	$F(15) = 6.06, p = 0.027^{*}$	$F(15) = 4.21, p = 0.058$	$F(15) = 8.45, p = 0.011^{*}$	$F(15) = 8.35, p = 0.011^{*}$
Parietal (P3, P4, Pz)	$F(15) = 2.00, p = 0.178$	$F(15) = 8.61, p = 0.010^{*}$	$F(15) = 2.43, p = 0.140$	$F(15) = 6.06, p = 0.027^{*}$	$F(15) = 10.33, p = 0.006^{*}$
Occipital (O1, O2)	$F(15) = 10.63, p = 0.005^{**}$	$F(15) = 9.44, p = 0.008^{*}$	$F(15) = 4.12, p = 0.061$	$F(15) = 5.86, p = 0.029^{*}$	$F(15) = 10.41, p = 0.006^{*}$
Temporal (T5, T6)	$F(15) = 8.37, p = 0.011^{*}$	$F(15) = 8.30, p = 0.011^{*}$	$F(15) = 2.61, p = 0.127$	$F(15) = 4.96, p = 0.042^{*}$	$F(15) = 11.98, p = 0.004^{**}$

Note: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

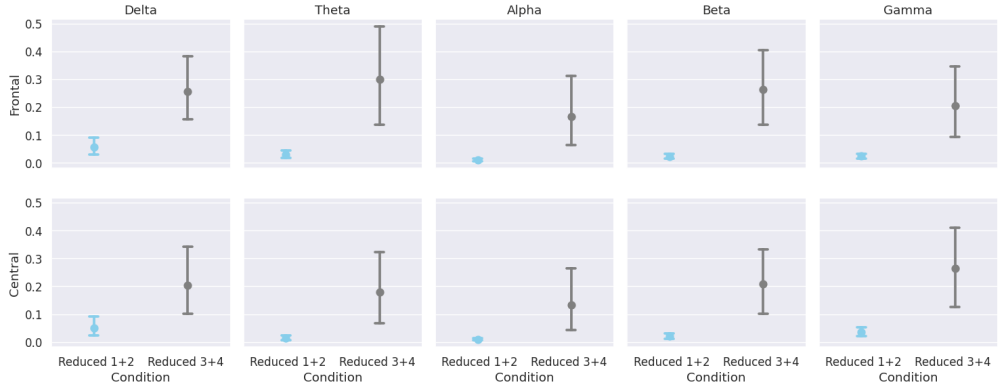


Fig. 4: Comparison of normalized PSD values between conditions negative reduced 1+2 and 3+4. The rows represent different brain regions, while the columns correspond to various frequency bands. This layout illustrates the directional differences in power across the specified conditions.

differences were primarily observed in the frontal and central brain regions. Consequently, we selected these regions for channel selection and trained a new model to further enhance classification performance.

TABLE IV: F1-Score Macro Results for All Channels (Negative Reduction 1+2 vs. 3+4)

Model	Delta	Theta	Alpha	Beta	Gamma
SVM	42.86	64.07	62.68	69.79	42.86
Neural Network	40.02	64.20	53.99	72.09	64.80
Neural Network with MMD Loss	42.65	64.53	57.58	**74.58**	67.39

As shown in Table V, the beta frequency band again achieves the highest F1-score across all three models. Notably, focusing on the frontal and central regions further improves performance, with the Neural Network with MMD Loss achieving an even higher F1-score of **75.57%** compared to 74.58%. This demonstrates that channel selection based on our statistical analysis enhances classification accuracy while reducing the number of EEG channels from 13 to 6.

TABLE V: F1-Score Macro Results for Frontal and Central Regions (Negative Reduction 1+2 vs. 3+4)

Model	Delta	Theta	Alpha	Beta	Gamma
SVM	42.11	63.64	55.42	66.62	62.73
Neural Network	53.58	63.18	53.99	67.88	62.20
Neural Network with MMD Loss	43.96	68.93	57.58	**75.57**	63.80

V. Discussion

This study investigated the neural mechanisms underlying ER by analyzing EEG data through frequency-domain and time-domain approaches and developing predictive models to classify successful versus unsuccessful ER.

Manuscript 1288 submitted to 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) . Received February 7, 2025.

Frequency-domain analysis revealed significant differences in PSD, particularly in the beta and gamma bands within the frontal and central brain regions, highlighting their crucial role in cognitive control and emotion modulation. These findings align with previous research demonstrating positive correlations between interoceptive awareness, mindfulness, and high-frequency EEG bands (beta, alpha, gamma), suggesting that ER involves similar high-level cognitive processes [22].

In predictive modeling, the Neural Network enhanced with MMD loss outperformed both the Support Vector Machine (SVM) and the baseline Neural Network, particularly after focusing on frontal and central channels. This underscores the effectiveness of MMD loss in addressing domain shifts and class imbalances in EEG-based applications. These results suggest that targeted machine learning techniques, combined with strategic channel selection, can effectively predict ER outcomes, offering potential applications in mental health interventions.

Despite its promising findings, the study has limitations. The relatively small sample size and imbalanced class distribution may constrain the generalizability of the results. Additionally, the exclusive focus on healthy young adults may limit the applicability of findings, as this demographic may be inherently better at ER. To address this, future research should consider augmenting datasets by including older adults or individuals with ER disorders, who may experience greater difficulty in regulating emotions. This could improve model robustness by introducing more variability in ER success and reducing the inherent bias from using only healthy participants. Future research should also aim to utilize larger and more diverse datasets, such as those involving video stimuli. This would allow for more realistic emotion evocation and help investigate ER in dynamic stimuli. Furthermore, optimizing the weighting factor for MMD loss and incorporating objective measures of emotional states could enhance model performance. Finally, additional domain adaptation methods should be employed to enhance the robustness and generalizability of predictive models.

VI. Conclusion

This study underscores the pivotal role of frontal and central brain regions and the beta frequency band in ER, as demonstrated through frequency-domain and time-domain analyses. The application of advanced machine learning models, particularly Neural Networks augmented with MMD loss, demonstrated robust predictive capabilities in classifying ER success. These findings highlight the potential of EEG-based biomarkers in the assessment of ER for personalized health interventions. To our knowledge, this integrated EEG-based approach to ER assessment is the first of its kind.

References

[1] K. McRae and J. J. Gross, "Emotion regulation," *Emotion (Washington, D.C.)*, vol. 20, no. 1, pp. 1–9, 2020.

[2] P. N. Lopes, P. Salovey, S. Côté, M. Beers, and R. E. Petty, "Emotion regulation abilities and the quality of social interaction." *Emotion*, vol. 5, no. 1, p. 113, 2005.

[3] H. J. Rosen, M. L. Gorno-Tempini, W. Goldman, R. Perry, N. Schuff, M. Weiner, R. Feiwell, J. Kramer, and B. L. Miller, "Patterns of brain atrophy in frontotemporal dementia and semantic dementia," *Neurology*, vol. 58, no. 2, pp. 198–208, 2002.

[4] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162–175, 2015.

[5] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2013, pp. 81–84.

[6] W.-B. Jiang, X.-H. Liu, W.-L. Zheng, and B.-L. Lu, "Seed-vii: A multimodal dataset of six basic emotions with continuous labels for emotion recognition," *IEEE Transactions on Affective Computing*, pp. 1–16, 2024.

[7] W. Liu, J.-L. Qiu, W.-L. Zheng, and B.-L. Lu, "Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 715–729, 2021.

[8] Z. Yin, M. Zhao, Y. Wang, J. Yang, and J. Zhang, "Recognition of emotions using multimodal physiological signals and an ensemble deep learning model," *Computer methods and programs in biomedicine*, vol. 140, pp. 93–110, 2017.

[9] P. Zhong, D. Wang, and C. Miao, "Eeg-based emotion recognition using regularized graph neural networks," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1290–1301, 2020.

[10] X. Lin, J. Chen, W. Ma, W. Tang, and Y. Wang, "Eeg emotion recognition using improved graph neural network with channel selection," *Computer Methods and Programs in Biomedicine*, vol. 231, p. 107380, 2023.

[11] V. Kosonogov, I. Ntoumanis, G. Hajiyeva, and I. Jääskeläinen, "The role of engagement and arousal in emotion regulation: an eeg study," *Experimental Brain Research*, vol. 242, no. 1, pp. 179–193, 2024.

[12] F. Del Popolo Cristaldi, G. Mento, G. Buodo, and M. Sarlo, "Emotion regulation strategies differentially modulate neural activity across affective prediction stages: An hd-eeg investigation," *Frontiers in Behavioral Neuroscience*, vol. 16, p. 947063, 2022.

[13] R. Thiruchselvam, J. Blechert, G. Sheppes, A. Rydstrom, and J. J. Gross, "The temporal dynamics of emotion regulation: An eeg study of distraction and reappraisal," *Biological psychology*, vol. 87, no. 1, pp. 84–92, 2011.

[14] T. A. Dennis and B. Solomon, "Frontal eeg and emotion regulation: Electrocortical activity in response to emotional film clips is associated with reduced mood induction and attention interference effects," *Biological psychology*, vol. 85, no. 3, pp. 456–464, 2010.

[15] Y. Song, Q. Zheng, B. Liu, and X. Gao, "Eeg conformer: Convolutional transformer for eeg decoding and visualization," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 710–719, 2022.

[16] W.-B. Jiang, X.-H. Liu, W.-L. Zheng, and B.-L. Lu, "Seed-vii: A multimodal dataset of six basic emotions with continuous labels for emotion recognition," *IEEE Transactions on Affective Computing*, 2024.

[17] W. Zheng, W. Liu, Y. Lu, B. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," *IEEE Transactions on Cybernetics*, pp. 1–13, 2018.

[18] S. Y. Dharia, M. Hojjati, S. G. Camorlinga, S. D. Smith, and A. S. Desroches, "Leveraging machine learning and threshold-free cluster enhancement to unravel perception of emotion and implied movement," in *IEEE-EMBS International Conference on Biomedical and Health Informatics*, 2024.

[19] C. E. MacKay, A. S. Desroches, and S. D. Smith, "An event-related potential (erp) examination of the neural responses to emotional and movement-related images," *Cognitive Neuroscience*, vol. 15, no. 1, pp. 1–11, 2024.

[20] M. Jas, D. A. Engemann, Y. Bekhti, F. Raimondo, and A. Gramfort, "Autoreject: Automated artifact rejection for meg and eeg data," *NeuroImage*, vol. 159, pp. 417–429, 2017.

[21] M. M. Murray, D. Brunet, and C. M. Michel, "Topographic erp analyses: A step-by-step tutorial review," *Brain Topography*, vol. 20, no. 4, pp. 249–264, 2008.

[22] R. A. Allegretta, K. Rovelli, and M. Balconi, "The role of emotion regulation and awareness in psychosocial stress: An eeg-psychometric correlational study," in *Healthcare*, vol. 12, no. 15. MDPI, 2024, p. 1491.