Offline EEG Signal Analysis to Enhance Visual Attention Training Platforms

Maryam Norouzi

***Abstract*—Attention is the ability to selectively process perceptually salient information while ignoring distractions that are irrelevant to an ongoing task. Visual attention, in particular, is the complex process of seeking out a target while filtering out competing stimuli. This study introduces a novel Brain-Computer Interface (BCI) platform designed to decode brainwave patterns associated with sustained attention in participants. We employed scalp electroencephalography (EEG) signals collection using a wireless headset during a visual attention task, where participants were primed to discriminate between composite images superimposed with scenes and faces, responding only to the relevant subcategory (e.g., indoor scenes) while ignoring the irrelevant ones (e.g., outdoor scenes). To understand the underlying neural mechanisms, we conducted Event-Related Potential (ERP) and wavelet transform analyses on the data. The ERP analysis revealed distinct neural responses to faces and scenes, with notable components such as N200, P300, and N600, suggesting differential engagement with these stimuli. The wavelet transform provided a time-frequency perspective, confirming the temporal dynamics observed in the ERP results and showing consistent patterns of neural activation across various EEG channels. Using deep learning techniques, we developed and optimized an individualized model to decode the attentional state of the participants based on their brainwaves. Our approach revealed instantaneous attention allocation towards the face and scene categories. Experiments conducted with three volunteer participants yielded an average decoding accuracy of approximately 85%, demonstrating the efficacy of our models. This work suggests potential applications for the evaluation of visual attention and the development of closed-loop brainwave regulation systems in the future.**

***Index Terms*—Visual Sustained Attention, BCI, EEG, Temporal/Spectral analysis, MLP classifier optimization.**

I. Introduction

Attention, a fundamental aspect of human cognition and perception, plays a crucial role in maintaining focus during tasks [1]. Deficits in attention are prevalent in various brain disorders like Alzheimer's disease, Traumatic Brain Injuries, and Posttraumatic Stress Disorder, where enhancing attentional states could improve cognitive functions like working memory [2]. Traditionally, attentional states have been studied using blood-oxygen-level dependent (BOLD) signals via functional magnetic resonance imaging (fMRI) [3], which, despite its high spatial resolution, has limitations for real-time neurofeedback due to the slow vascular response of the brain [4]. Electroencephalography (EEG), in contrast, offers a more direct and timely measure of neural activity, making it suitable for real-time brain-computer interface (BCI) applications [5]. EEG has been widely used for attention evaluation and training, with significant progress in neurofeedback training for attention disorders, particularly in children with Attention Deficit/Hyperactivity Disorder [6]. Since face-like visual stimuli undergo specialized processing in the human brain compared to other non-face objects [7] face images were employed in numerous brain studies particularly studies on attention [8]. In recent offline work, List et al. [9] explored the spatiotemporal changes in EEG to classify perceptual states (e.g., faces versus Gabors) and analyzed the scope of attention (e.g., locally versus globally focused states). Sreenivasan et al. [10] used event-related potential (ERP) analysis to show that attention to faces in composite images with different transparency of face images can modulate perceptual processing of faces at multiple stages of processing. Nevertheless, this study did not report the EEG classifcation results when the subjects attended to both categories of images in the overlapped images in separate blocks of training. We hypothesize that scalp EEG data collected from an individual who attends to the elements of a complex stimulus contains relevant and distinguishable patterns. The majority of previous studies focused on identifying the level of attention in a participant without knowing which visual stimulus the sustained attention is devoted to.

This study introduces a BCI system that records EEG signals during a visual attention task featuring composite images of faces and scenes. ERP and wavelet transform analyses are then applied to identify unique brain wave patterns in response to these stimuli. A multi-layer perceptron (MLP) classifier is subsequently developed, optimized, and personalized for each participant.

II. Materials and methods

This section outlines the components and materials used in the Brain-Computer Interface (BCI) system. It describes the procedure for collecting scalp EEG data from participants during a sustained attention task and details of the techniques used for EEG data analysis.

A. Development of the BCI Platform

The BCI system includes a wireless EEG headset, a workstation computer equipped with dual monitors (one for human subjects and one for the practitioner), and a developed Python-based software for data acquisition and analysis. Figure 1 shows the platform's components and the flow of data between them. As well, a Graphic User Interface (GUI) has been created to enable easy management of the experimental protocol by a practitioner, Figure 2.

1) EEG Recording Device: EEG signals were acquired using a wireless headset called Unicorn Hybrid Black. The headset has 8 channels of EEG electrodes located based on 10-20 international systems covering the central, parietal,

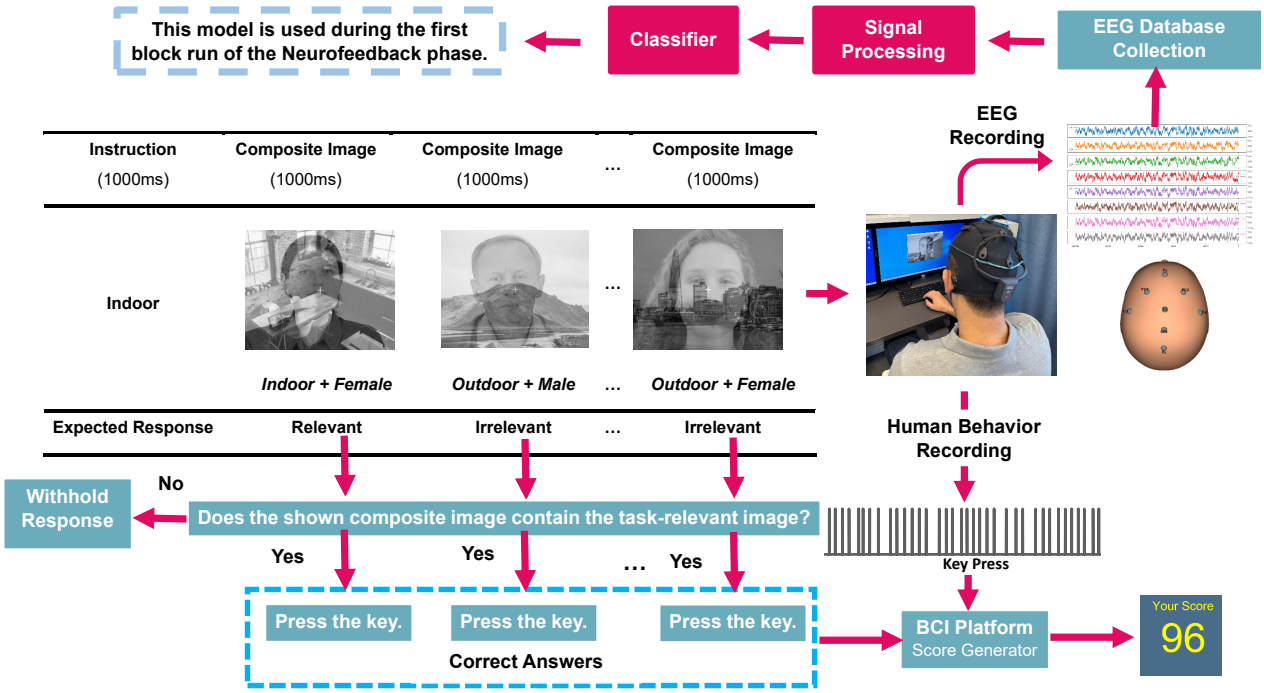


Fig. 1: Developed BCI Platform for neurofeedback visual attention training, (clinical trial number: NCT05908253). The pre/post-evaluation (offline/ open loop) phase has eight blocks, starting with a 5-second text cue, followed by 7-second baseline images (5-second grey, 2-second black with cross sign), and 40-second activity images (40 1-second trial images at 50% transparency). Human responses are recorded, and an attention score is computed at the end of this phase, also, utilizing acquired EEG data, the initial classifier is developed and used for the first block run of the neurofeedback phase.

Fig. 2: Graphical User Interface of the EEG-Based BCI System for Visual Attention Assessment

occipital, and part of the frontal regions of the scalp. The exact locations are labeled sequentially as Fz, C3, Cz, C4, Pz, Po7, Oz, Po8. The sampling frequency was set to 250 Hz. The device collects brain signals and transmits the participant's brain signals to the PC with a Bluetooth connection. The use of wireless headsets offers the potential for applications in real-life settings [11].

2) Interface: We performed our data acquisition and analysis in Python. The visual stimulation and designed protocols were also controlled by Python through a customized Graphic User Interface (GUI). Upon executing the code, a GUI interface will appear. Below are the detailed steps for using the GUI:

- Input the patient's name and click the 'Create Folder' button. This action will generate a main folder containing three subfolders, each corresponding to one of the three phases, where the collected data will be stored.
- Select the desired phase—namely, Pre-evaluation, Neurofeedback, or Post-evaluation. For the scope of this project, the focus will be exclusively on the Pre-evaluation phase.
- Click the 'Create Trial' button to generate a random order of blocks for each run.
- Press the "Start Trial" button.
- Monitor the progress by viewing the "Last Block Completed" widget and the "Phase Progress" widget. After the completion of each block, a 10-second rest period follows before the subsequent block begins. This sequence repeats until the progress bar is filled.

3) Stimuli: Four subcategories of indoor scenes, outdoor scenes, male face, and female face images were chosen in our experiment as stimuli. The images were all black and white with equal sizes, 800 × 1000 pixels. Face images were chosen to be neutral without any emotional expression. They were centered inside the composite image. The indoor images were chosen from interior scenes. Outdoor images were natural landscapes and cityscapes. Brightness and contrast for all face and scene images were adjusted so that the images on both face and scene categories have equalized and identical contrast. Participants were trained with a sequence of superimposed images of face and scene. The duration of each stimulus was set to be 1000 milliseconds. It was programmed to have 50% transparency of each image subcategory in the superimposed

images. The platform has also the capability to further adjust the transparency of image types in the superimposed image based on the attention level of the instructed subcategory. In the middle of the images there is a cross sign in which subjects are asked to fixate their gaze on the middle of the screen.

B. Experimental Paradigm

Three healthy participants including 2 males and 1 female, with a mean age of 29 years, voluntarily completed eight training blocks of the experiment. All participants had normal or corrected-to-normal vision. All participants gave written consent to the experiment. The computerized task was provided by a PC with dual monitors. One monitor was viewed by the experimenter to control the experiment. The other monitor was positioned in front of the participants for the presentation of stimuli. The participants were asked to sit comfortably in a fixed chair with one hand resting on the lap and another hand ready to press a key to give behavioral responses. The participants were instructed to pay attention to the monitor during the experiment and limit their excessive body movement. The participants were also asked to fixate their gaze on the middle of the screen and keep their heads at approximately 50 cm from the monitor while observing the stream of images. Our experimental protocol consisted of eight blocks of trials with a 10-second respite between blocks. The experiment starts with 3 minutes just recording to make sure that the EEG signals are settled, followed by eight blocks of trials Figure 3. Each block started with a five-second texture cue instructing the attended subcategory image, followed by 5 seconds of a grey image, 2 seconds of a black image with a cross sign in the middle of the image (considered as baseline data) followed by 40 trials of image stimuli. The duration of each trial was set to one second without any intertrial time. A trial includes a grey scale overlaid picture in which 50% of opacity was from the scene (indoor or outdoor) category and 50% was from the face (male or female) category. There was no repetition of face or scene images through each block of the experiment. This process helped to prevent any learning mechanism from happening for the participant. Participants were asked to identify whether the shown image contained the task-relevant image (e.g., an indoor image) or the task-irrelevant image (e.g., an outdoor image) by responding to each superimposed image. They were asked to press the key on the keyboard for each recognized relevant image and withhold their responses for irrelevant images. 90% of the composite images contained images from the task-relevant subcategory (e.g., indoor image) while the other 10% of the composite images contained images from the task-irrelevant subcategory (e.g., outdoor image). Figure 1 illustrates a sample sequence of composite images during a block and also the corresponding expected responses from participants. we ran the experiment three times to get more trials. The total duration of the experiment was approximately 60 minutes per participant, inclusive of the time spent donning the EEG cap, checking electrode impedance, and taking rest between runs.

C. Signal processing

In the present work, we aimed to identify participants' attentional states into two categories of images (face versus scene; regardless of their subcategories) by using recorded EEG signals. The participants were primed with the subcategories throughout the experiment. So, we hypothesized that the brainwaves contained common features for the subcategories of one category. This assumption reduced the problem of the classification of EEG signals to a 2-class classification problem, i.e., classifying underlying patterns of EEG while the participants attended to faces or scenes. Flowcharts in Figures 1,3 illustrate the process of analyzing a participant's overt response as well as his/her EEG signals. A brief description of EEG signal preprocessing, temporal and spectral analysis, and classifcation techniques is given as follows.

1) Signal pre-processing: The pre-processing of the EEG data involved several key steps to ensure signal quality and accuracy:

1) **Bandpass Filtering:** Each EEG channel was subjected to a 5th-order Butterworth bandpass filter, which was used to retain frequencies between 0.4 Hz and 40 Hz. This step was crucial for reducing low-frequency drift and high-frequency noise, ensuring that only relevant frequency components of the EEG signals were retained.
2) **Artifact Rejection:** To mitigate artifacts in the EEG signal, a technique involving the detection of spikes based on the Median Absolute Deviation (MAD) was employed. Once these spikes were identified, they were interpolated using a Cubic Spline, thereby preserving the integrity of the EEG data.
3) **Denoising:** A K-Neighbors Regressor was utilized for denoising each EEG channel. This method predicted the signal based on surrounding time points, effectively smoothing the EEG data and reducing random noise.
4) **Epoching:** The continuous EEG signal was divided into smaller segments, referred to as epochs, based on the sampling rate. This process allowed for the isolation of specific events or stimuli in the EEG recording for detailed analysis.
5) **Spike Detection:** Abnormal spikes were detected in each epoch and channel. Epochs containing these spikes were marked for exclusion to prevent the contamination of the dataset with large artifacts, which could significantly skew the results.
6) **Epoch Concatenation:** Finally, epochs that were free of artifacts were concatenated back into a continuous dataset. This reconstructed dataset was then used for further analysis, ensuring that the data was clean and representative of the true EEG signal.

These pre-processing steps established a solid foundation for precise EEG signal analysis, thus facilitating a more dependable interpretation of brainwave data. Following these steps, the spike-rejected data is transformed into a format compatible with the MNE-Python library, a powerful tool for neurophysiological data analysis, including Independent Component Analysis (ICA). ICA separates a multivariate signal into independent non-Gaussian components. For observed

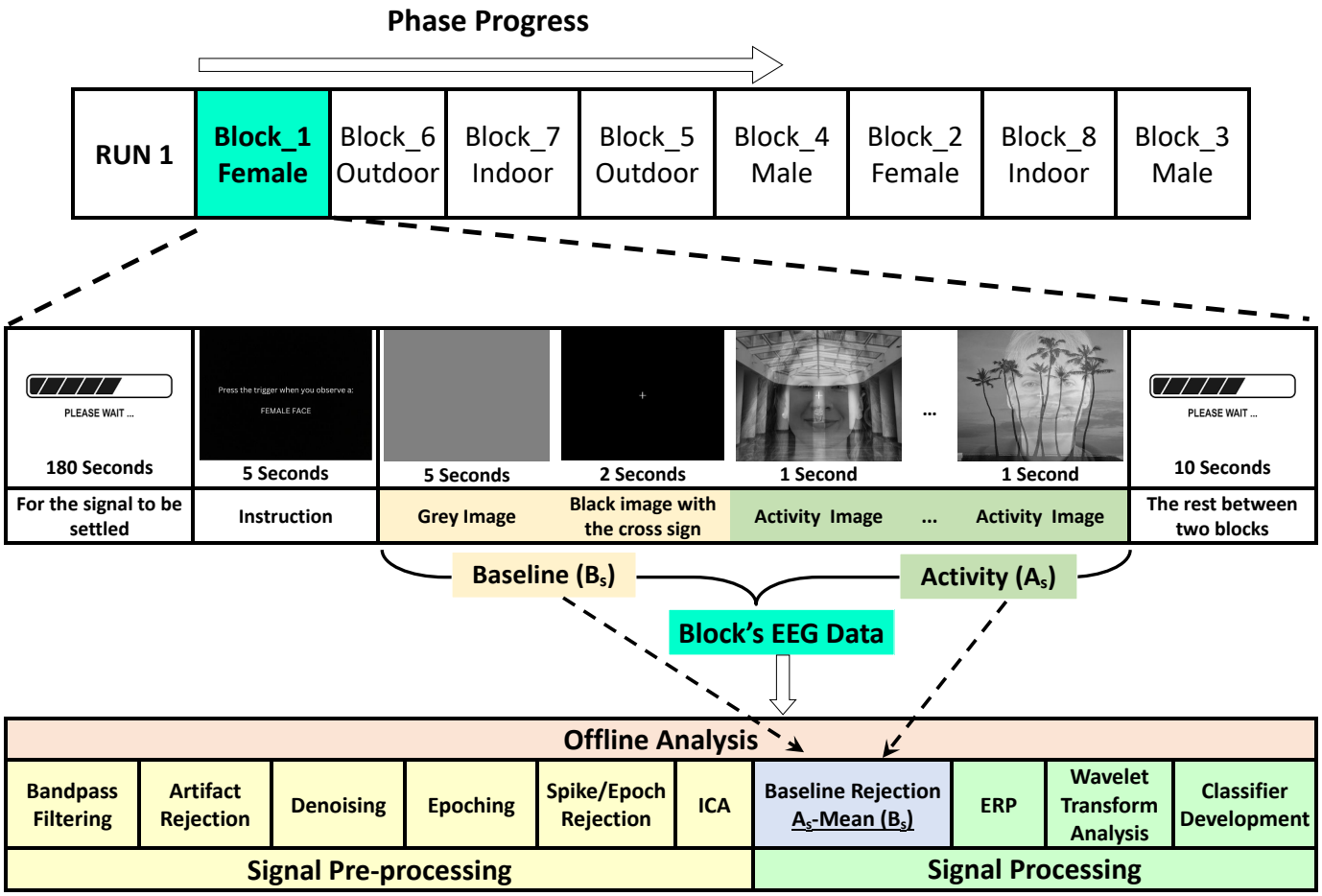


Fig. 3: Flowchart of EEG experiment and data analysis process: showing one run out of three per participant, including eight randomized blocks of trials. The run starts with a 3-minute EEG signal settling phase, followed by a 5-second instruction display, a 5-second grey image, and a 2-second black image with a cross. This precedes 40 one-second activity images, followed by 10-second rests between blocks. The recorded data is then preprocessed and analyzed to develop a binary classifier.

mixed signals $\mathbf{X}$, ICA computes a separating matrix $\mathbf{W}$, so that $\mathbf{S} = \mathbf{WX}$ is a matrix of independent components. In this project, we employed Infomax ICA, a method that maximizes mutual information between the input and output of a neural network model. This approach effectively isolates and removes artifact components, enhancing signal clarity. Additionally, the ICA procedure is executed cyclically throughout the study, allowing for continuous refinement of the data. It commences with the visualization of each block's ICA components, which aids in the identification and subsequent iterative removal of artifacts such as eye blinks or muscle movements. Following the ICA, there is a phase of reviewing and approving the improved signals. Should the results not meet the desired standards, the option to reapply the ICA process is available. The final step involves the user's confirmation of the data's integrity, post which the purified signals are preserved for additional analysis. This process is then reiterated for subsequent blocks of data, ensuring a thorough and meticulous approach to data cleansing.

D. Signal Processing

In the present work, we aimed to identify participants' attentional states into two categories of images (face versus scene; regardless of their subcategories) by using recorded EEG signals. An overview of EEG signal analysis in both time and time-frequency domains, along with the development of classifiers, is provided below.

1) Time-Domain Analysis: Event-related potentials (ERPs) are categorized under time-domain analysis in brain signal processing. ERPs are measured responses in the brain that are the direct result of a specific sensory, cognitive, or motor event. Unlike wavelet analysis, which decomposes the signal into frequency and time, ERP analysis focuses on the amplitude and latency of the brain's electrical activity in response to events, captured over time. The analysis is concerned with averaging the EEG signal over several trials to identify voltage fluctuations associated with specific events, thus enhancing the signal-to-noise ratio and isolating the specific brain responses from the ongoing EEG. After the initial data pre-processing, the segmentation of the EEG recordings of each block was executed, wherein each segment of 250 samples—equivalent to one second of recording or 1000 milliseconds—was con-

sidered as an epoch. These epochs were divided into two categories: the baseline set, consisting of epochs aligned with the displaying of neutral stimuli (black/grey images), preceding the experimental trials, and the activity set, corresponding to epochs linked with the composite stimulus images. To increase the signal-to-noise ratio, we did the baseline rejection, the mean voltage level of the baseline epochs was computed and subsequently subtracted from each data point within the experimental epochs to control for pre-stimulus voltage fluctuations. This step makes sure that the brain's reaction is measured from a consistent starting point, highlighting any changes caused by the stimulus. Temporal resolution reduction was then applied to the epochs by averaging every five consecutive samples, thereby downsampling the data from 250 to 50 samples per epoch. This downsampling serves to enhance the signal-to-noise ratio and reduce the computational load without compromising the temporal dynamics of the ERPs. Each epoch category was then aggregated to obtain two distinct mean signals per block, corresponding to the baseline and stimulus-evoked neural activities, respectively. This process was consistently applied across all 24 blocks of the study. The next step in the analysis was to separate the data into face and scene groups, and then average them to get an overall brain activity pattern for each category of image. The culmination of this analysis is presented as distinct ERP waveforms for face and scene stimuli, offering a visual representation of the temporal neural responses, as depicted in Figure 5.

2) Time-Frequency Analysis: In our study, the Morlet wavelet transform was employed for the analysis of EEG data. This process began with the generation of a Morlet wavelet for each frequency within the predefined range (0 to 40 Hz, in 2 Hz steps). The number of cycles for each wavelet varied linearly from 1 to 10, enabling a balanced time-frequency resolution tailored to each frequency. Each Morlet wavelet was then convolved with the EEG data. This convolution process isolated the frequency-specific components of the EEG signals, resulting in a time-series representation of the power for each frequency. Following the Morlet wavelet transformation of EEG data, we proceeded with epoching, which involves dividing the continuous EEG signal into shorter, fixed-length segments or epochs. Each epoch consisted of 250 data points, equivalent to one-second intervals given our sampling rate, consistent with the methodology applied in ERP analysis. Crucially, within each data block, epochs were categorized into two distinct sets: 'base' and 'activity.' The 'base' epochs were identified as periods representing baseline brain activity. Conversely, 'activity' epochs corresponded to periods where specific events or stimuli were presented. This separation was essential for subsequent baseline normalization and analysis of stimulus-specific brain activity. For each block, the mean power of all 'base' epochs was computed, resulting in a single, average baseline signal. Similarly, an average signal for the 'activity' epochs was obtained by calculating the mean power across these epochs. This process effectively distilled the complex EEG data into two representative signals per block, one reflecting baseline brain activity and the other depicting brain responses to stimuli or events. In the crucial step of baseline normalization, we focused specifically on the first 200 data points of the mean 'base' signal. This subset of the baseline was used to normalize the 'activity' signal, thereby controlling for variations in baseline brain activity across trials and conditions. The normalization involved dividing the power values in the 'activity' signal by the mean of these first 200 baseline data points. This targeted approach to normalization ensured that the resultant activity epochs were adjusted relative to a consistent baseline measure, enhancing the reliability and comparability of our findings across different blocks and conditions. The normalization was quantified using the following logarithmic ratio:

$$dB = 10 \cdot \log_{10} \left(\frac{\text{activity}}{\text{baseline}} \right)$$

Finally, we calculated the mean power across all blocks to produce a signal representative of the neural response to the stimuli across all recorded sessions.

3) Classifier Development: In this study, the development of a MLP c ssifier is approached through a meticulous optimization process utilizing Optuna, a hyperparameter optimization framework. MLP is a type of feedforward neural network defined by the following equations:

- For the input layer:

$$\mathbf{x}_0 = \mathbf{x}_{\text{input}} \tag{1}$$

- For hidden layers ($l = 1, \ldots, L$), where L is the number of hidden layers:

$$\mathbf{x}_l = f(\mathbf{W}_l \mathbf{x}_{l-1} + \mathbf{b}_l) \tag{2}$$

  Here, $\mathbf{x}_l$ is the output of layer l, $\mathbf{W}_l$ is the weight matrix, $\mathbf{b}_l$ is the bias vector, and f is the activation function which can be *relu*, *logistic*, *tanh*, or *identity*.
- For the output layer:

$$\mathbf{y}_{\text{output}} = \mathbf{W}_{\text{output}} \mathbf{x}_L + \mathbf{b}_{\text{output}} \tag{3}$$

  where $\mathbf{y}_{\text{output}}$ is the final output of the network.

The MLP is trained to minimize a loss function, typically cross-entropy for classification tasks, using backpropagation and gradient descent (or its variants). The network parameters ($\mathbf{W}_l, \mathbf{b}_l$) are optimized through iterative training. The initial learning rate and the number of iterations are crucial hyperparameters in this process. The objective of this project is to find the most effective combination of hyperparameters that would enhance the performance of the MLP model on a given dataset for each subject. The key hyperparameters under optimization include the number of layers (ranging from 1 to 3), the number of neurons in each layer (between 16 to 512), the type of activation function (options include: 'relu', 'logistic', 'tanh', 'identity'), the initial learning rate (ranging logarithmically from 1×10^{-4} to 1×10^{-1}), and the maximum number of iterations for the solver (between 50 and 1000). The optimization process employs Stratified K-Fold cross-validation with five folds to evaluate the performance of various hyperparameter configurations. This method ensures a robust assessment by training and validating the model on different subsets of the data, maintaining the original proportion of classes in each

fold. The performance of each configuration is measured based on the accuracy achieved on the validation sets across all folds. The average accuracy across these folds serves as the primary metric for Optuna to identify the best set of hyperparameters. Once the best hyperparameters are determined, the final MLP model is configured accordingly and trained on the entire training dataset. This step leverages the insights gained from the cross-validation phase to capitalize on the full training data, thus refining the model's ability to generalize. Subsequently, the fully trained model's performance is evaluated on an untouched test dataset (20% of data), which serves as a critical assessment of the model's predictive capability on new, unseen data. This final test accuracy offers a transparent and unbiased evaluation of the model's effectiveness, ensuring that the final MLP classifier is both robust and reliable for the task at hand.

III. Results and discussion

A. ICA-Based Artifact Reduction

The application of ICA on a single block of EEG data for one subject, as shown in Figure 4, has successfully decomposed the signal into multiple components, each representing distinct sources of activity. Component 1 stands out with a broad, high-frequency spectrum in its power spectral density (PSD), indicative of muscle artifact contamination. Given its characteristic non-brain artifact profile, Component 1 has been appropriately removed to minimize interference with the neural signals. The remaining components (2-7) present mixed signals with both brain-derived activity and potential artifacts. For instance, Components 2, 3, and 5 exhibit features that align with typical EEG rhythms, such as the alpha and beta waves, suggesting their cerebral origin. However, they also contain traces of artifacts, which could be eye movements or other physiological noises. Despite this, they have been retained to avoid losing valuable brain activity information embedded within these components. Components 4, 6, and 7, while also displaying some artifact characteristics, show spatial distributions and time courses consistent with EEG signals. For example, Component 4 might contain some eye-blink artifacts, yet its frequency profile also captures relevant EEG dynamics. Similarly, Components 6 and 7, despite the possibility of containing electrode noise or minor muscle artifacts, primarily represent neural activity. The decision to retain these components is a balance between artifact removal and the preservation of neural data. Excessive elimination of components could lead to a loss of critical EEG information, which might be more detrimental than the minor artifacts that they carry. The data indicates that Component 1, removed through the ICA process, was predominantly characterized by muscle artifact noise, particularly from the head and neck regions, affecting the C3 and PO7 channels. By discarding this component, the EEG data for these channels is now largely free from the distortions that such muscular artifacts cause, leading to a more accurate reflection of brain activity in the post-ICA signal representations, affirming the effectiveness of the ICA process in enhancing the quality of EEG data for further analysis.

B. ERP Analysis of Faces vs. Scenes

The ERP results depicted in Figure 5 provide insight into the neural processing of visual stimuli categorized as faces and scenes during a visual attention task, with a baseline measure included for contrast. The data exhibits characteristic waveforms associated with well-known ERP components such as the N200, P300, and N600, which vary across the different EEG channels and stimulus types. In the frontal channel (Fz), the scene stimulus elicits a distinct negative deflection around the 200 ms mark, likely corresponding to the N200 component. This component is often associated with conflict detection and cognitive control mechanisms, suggesting that scenes may be triggering more complex cognitive processes in this early stage compared to faces. The central channels (C3 and C4) display prominent positive peaks around 300 ms post-stimulus, which aligns with the P300 component. This component is related to attention and the processing of stimulus salience. The presence of P300 in response to both faces and scenes indicates engagement with and differentiation between these stimuli types but without a pronounced lateralization effect. At the parietal channel (Pz), a notable response is observed for the face stimulus around the 600 ms mark, which could be indicative of the N600 component. This late negative deflection is often linked to higher-level cognitive processes, such as memory recall and updating, and semantic processing. The N600 response to faces at Pz suggests that the neural processing of faces may involve more complex associative mechanisms compared to scenes, possibly due to the social and emotional relevance of faces. Additionally, lateralization effects are observed in the occipito-parietal channels (Po7 and Po8), with the channels showing differential responses to the two types of stimuli. This suggests that the neural processing of faces and scenes not only varies in magnitude but also in spatial distribution across the hemispheres, which is consistent with the known lateralization of visual processing in the brain. Overall, the ERPs across these channels indicate selective attentional processes and differentiated cognitive engagement with faces and scenes. The P300 findings across the scalp reflect a broad attentional allocation and working memory engagement with both stimuli types. In contrast, the N200 and N600 components suggest more complex and possibly stimulus-specific cognitive processes at work, with a degree of lateralization that underscores the distinct pathways for processing these visual categories.

C. Wavelet Transform Analysis

The wavelet analysis depicted in Figure 6 shows the time-frequency representation of the brain's response to face and scene stimuli across various EEG channels. This method reveals the power (in dB) across different frequencies over time, providing a more nuanced view of the brain's dynamic activity in response to visual stimuli. In the frontal channel (Fz), there's a notable increase in power for both stimuli types in the lower frequency bands (below 10 Hz), which may correspond to the delta and theta bands associated with attention and cognitive processing. There's a visible peak in power for scene stimuli around 200-400 ms, which could

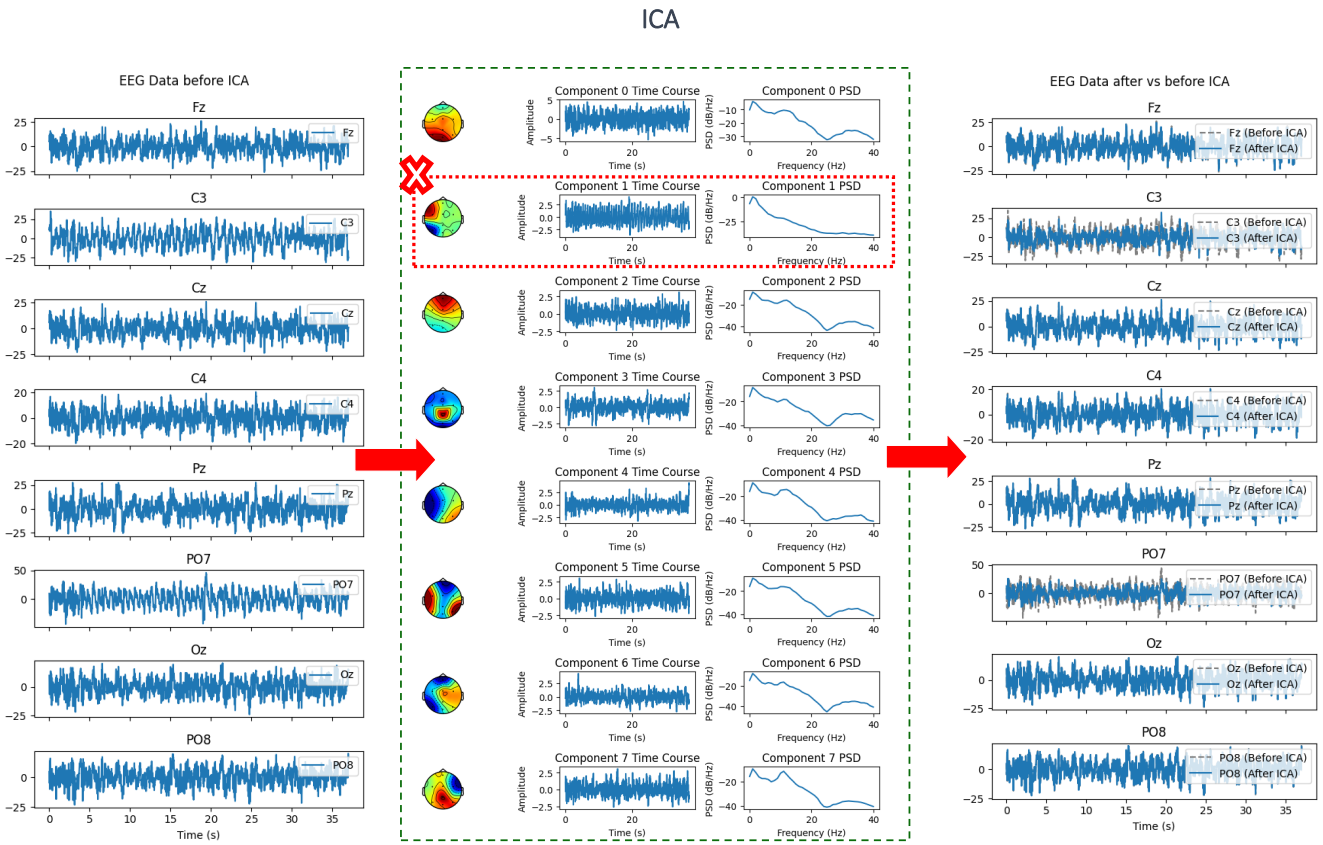


Fig. 4: Strategic Artifact Rejection in EEG via ICA: Targeted Exclusion of Component 1 and Its Effects on C3 and PO7 Channels. Component 1 likely represents muscle artifact contamination. This is inferred from its broad high-frequency spectrum in the PSD. The removal of such artifacts is critical, especially for channels like C3 and PO7, which are susceptible to muscle noise due to their proximity to the muscles of the neck and head. The ICA approach has effectively isolated and removed this noise, which is evident in the post-ICA data comparison. The cleaner signals in these channels post-ICA suggest successful artifact mitigation, likely leading to a more accurate representation of cerebral activity.

correlate with the N200 component observed in the ERP results, suggesting an attentional response to the visual scenes. In the central channels (C3, Cz, C4), there's an increase in power in the alpha and beta bands (around 10-30 Hz) for both types of stimuli. This is particularly evident in the Cz channel for scenes, which could correspond to the P300 component seen in the ERP analysis, indicating an engagement with the stimulus and attentional processing. The parietal channel (Pz) shows a strong increase in power for both stimuli in the alpha and beta bands. The consistency with the ERP results, where an N600 component was noted, suggests sustained cognitive processing, likely related to the categorization and integration of the stimuli. In the occipital channels (Oz, Po7, Po8), which are more directly involved in visual processing, there's a distinct pattern of power increase across a broad range of frequencies, most prominently for the scene stimuli. This might be indicative of the visual system's engagement in processing complex visual scenes, which aligns with the ERP findings of a more pronounced response to scenes in these regions. Comparing the wavelet analysis to the ERP results, there is a consistency in the temporal dynamics of the neural response to both faces and scenes. The ERP components such as N200, P300, and N600, which reflect different cognitive processes, are also observed as increases in power at corresponding times in the wavelet analysis. This cross-validation between methods reinforces the reliability of the observed neural responses to the visual stimuli and provides a comprehensive understanding of the temporal and spectral dynamics involved in visual attention tasks.

D. Individualized MLP Optimization for EEG-Based BC

The optimization of the Multilayer Perceptron (MLP) model was conducted separately for each of the three subjects, ensuring personalized parameter tuning without data intermixing. The outcomes, detailed in Table I and reflected in Figure 7, underscore a tailored approach to EEG-based Brain-Computer Interface (BCI) system design. Subject S1 achieved an accuracy of 83% utilizing a simplified single-layer model, which may suggest a less complex neural pattern or a dataset that can be adequately characterized with a less granular approach. For Subject S2, the introduction of additional layers, a total of four, alongside a ReLU activation function, culminating in an enhanced accuracy of 86%. This indicates that S2's data perhaps possessed more intricate features, which were better captured by the added complexity of the model. Subject

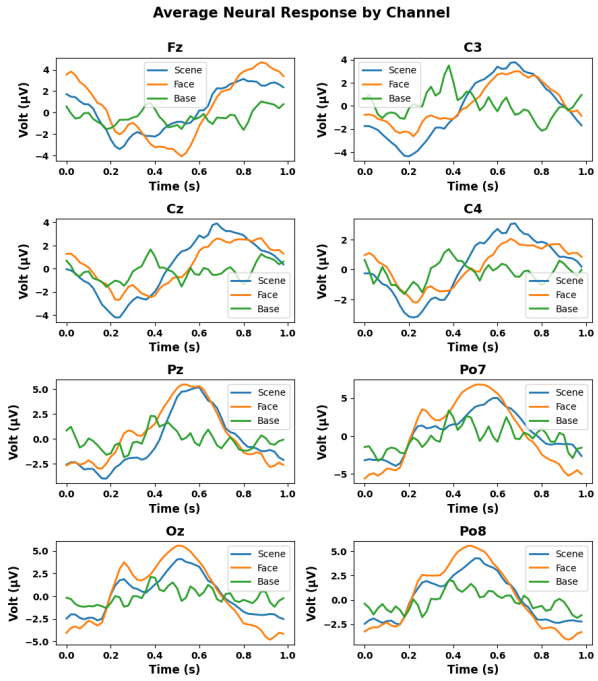


Fig. 5: Comparative ERP Responses to Facial and Scenic Stimuli Across EEG Channels: Display of N200, P300, N600 Components Related to Cognitive Attention and Stimulus Assessment.

S3's data yielded the highest accuracy at 88%, with a three-layer structure and a logistic activation function, suggesting an optimal balance between model complexity and the ability to capture relevant patterns in the EEG data for this particular subject. The differing architectures, neuron counts, and activation functions reflect the individualized nature of the EEG signals and the necessity for distinct model configurations. These results were validated using a 5-fold cross-validation approach, providing a rigorous assessment of the model's predictive power. The preserved 20% of data for testing ensures that the reported accuracies genuinely reflect the model's performance on new, unseen data. This individual-centric optimization strategy emphasizes the potential for customizing BCI systems to enhance performance and reliability for each user.

TABLE I: MLP optimization results across various subjects.

	S1	S2	S3
accuracy	0.83	0.86	0.88
n_layers	1	4	3
n_units_layer0	329	114	414
n_units_layer1	-	247	355
n_units_layer2	-	461	285
n_units_layer3	-	173	-
activation	tanh	Relu	logistic
learning rate	0.0032	0.0010	0.0103
max_iter	518	605	69

IV. Conclusion

In this study, we focused on the development and implementation of a Brain-Computer Interface (BCI) system

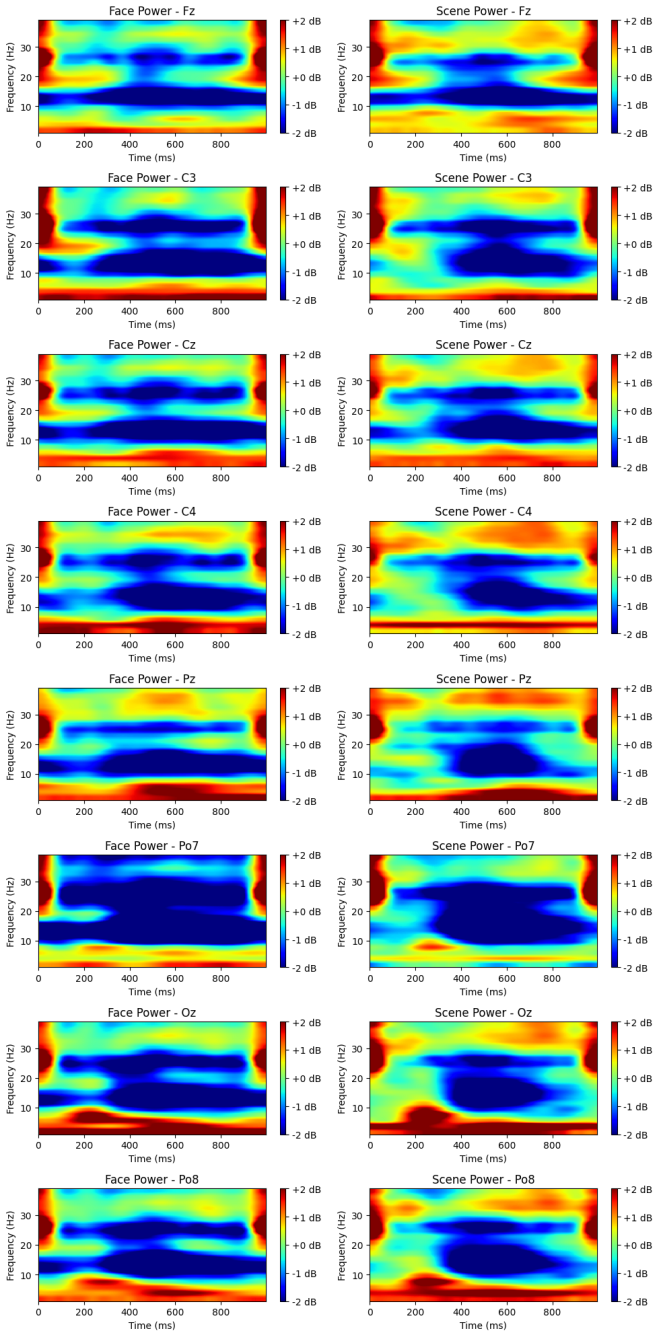


Fig. 6: Time-Frequency Analysis of Visual Stimuli: Face vs. Scene Power Spectra in EEG Channels. The time-frequency plots display the dynamic power spectrum of EEG responses to facial and scenic stimuli across different brain regions. The data shows distinctive patterns of neural activity, varying across frequencies and time points, which reflect the brain's processing of these different visual categories. Peaks in power within specific frequency bands suggest the differential engagement of cognitive resources, possibly linked to attentional and recognition processes activated by each stimulus type.

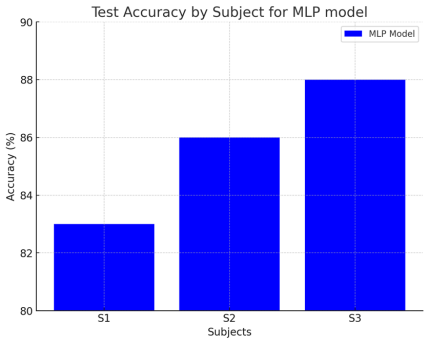


Fig. 7: Subject-Specific MLP Model Accuracy for BCI Systems.

for analyzing visual attention using EEG signals. This involved using a wireless EEG headset, workstation computers, and a novel Python-based software for data collection and analysis. The study participants were exposed to a series of visual stimuli while EEG data was collected. Signal processing techniques including bandpass filtering, artifact rejection, denoising, epoching, and Independent Component Analysis (ICA) were used to ensure signal quality. Results show the effectiveness of the BCI system in distinguishing between attentional states towards different image categories. The results showed distinct neural responses to faces and scenes, with significant components such as N200, P300, and N600 indicating differential engagement with stimuli. A time-frequency analysis using Morlet wavelet transform further confirmed these findings. An individualized Multi-Layer Perceptron (MLP) classifier was developed for each participant, achieving an average decoding accuracy of approximately 85%, demonstrating the efficacy of the tailored models.

From the results of this study, we can develop a novel closed-loop decoder adaptation pipeline for a neurofeedback-based attention training platform. An EEG-based real-time BCI system that simultaneously records brain signals and classifies brainwave patterns in milliseconds to provide instant and adjustable neurofeedback (visual stimuli: changing the transparency of the composite image based on the subject's visual attention performance) can be developed. The given neurofeedback stimulus is meant to encourage the human subject to modify his/her neural activity for maximum attention and feedback reward. Also, we can leverage this platform and develop an adaptive online decoder for closed-loop neurofeedback which could be initialized by a set of training data collected during an open-loop form.

REFERENCES

[1] Mehdi Ordikhani-Seyedlar, Mikhail A Lebedev, Helge BD Sorensen, and Sadasivan Puthusserypady. Neurofeedback therapy for enhancing visual attention: state-of-the-art and challenges. *Frontiers in neuroscience*, 10:352, 2016.

[2] Yang Jiang, Reza Abiri, and Xiaopeng Zhao. Tuning up the old brain with new tricks: attention training via neurofeedback. *Frontiers in aging neuroscience*, 9:52, 2017.

[3] Robert T Thibault, Michael Lifshitz, and Amir Raz. The self-regulating brain and neurofeedback: Experimental science and clinical promise. *cortex*, 74:247–261, 2016.

[4] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *sensors*, 12(2):1211–1279, 2012.

[5] Reza Abiri, Soheil Borhani, Eric W Sellers, Yang Jiang, and Xiaopeng Zhao. A comprehensive review of eeg-based brain–computer interface paradigms. *Journal of neural engineering*, 16(1):011001, 2019.

[6] Choon Guan Lim, Chui Pin Soh, Shernice Shi Yun Lim, Daniel Shuen Sheng Fung, Cuntai Guan, and Tih-Shih Lee. Home-based brain–computer interface attention training program for attention deficit hyperactivity disorder: A feasibility trial. *Child and Adolescent Psychiatry and Mental Health*, 17(1):1–11, 2023.

[7] Bruno Rossion, Corentin Jacques, and Jacques Jonas. Intracerebral electrophysiological recordings to understand the neural basis of human face recognition. *Brain Sciences*, 13(2):354, 2023.

[8] Greta Tuckute, Sofie Therese Hansen, Troels Wesenberg Kjaer, and Lars Kai Hansen. Real-time decoding of attentional states using closed-loop eeg neurofeedback. *Neural Computation*, 33(4):967–1004, 2021.

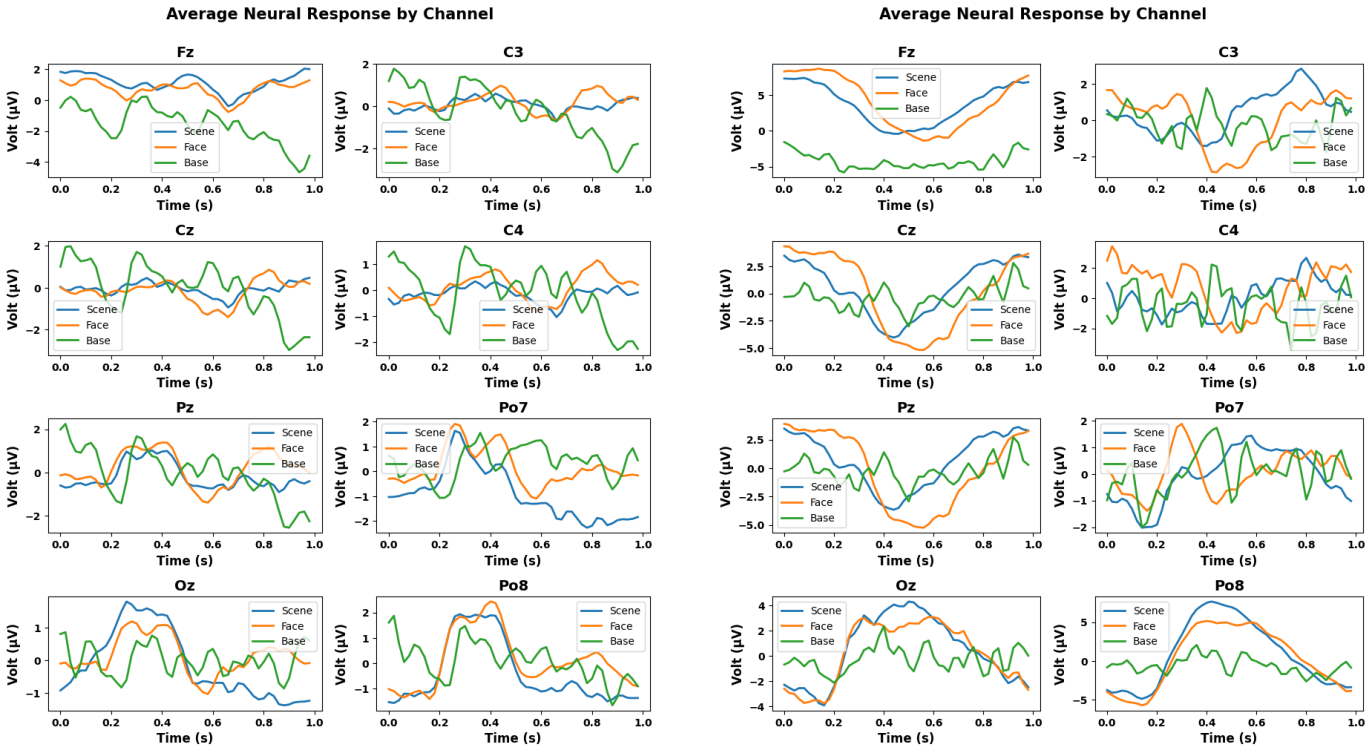
[9] Alexandra List, Monica D Rosenberg, Aleksandra Sherman, and Michael Esterman. Pattern classification of eeg signals reveals perceptual and attentional states. *PloS one*, 12(4):e0176349, 2017.

[10] Kartik K Sreenivasan, Jonathan M Goldstein, Audrey G Lustig, Luis R Rivas, and Amishi P Jha. Attention to faces modulates early face processing during low but not high face discriminability. *Attention, Perception, & Psychophysics*, 71:837–846, 2009.

[11] Rob Zink, Borbála Hunyadi, Sabine Van Huffel, and Maarten De Vos. Mobile eeg on the bike: disentangling attentional and physical contributions to auditory attention tasks. *Journal of neural engineering*, 13(4):046017, 2016.

APPENDIX A
COMPARATIVE ANALYSIS OF ERP AND TIME-FREQUENCY RESPONSES ACROSS HUMAN SUBJECTS

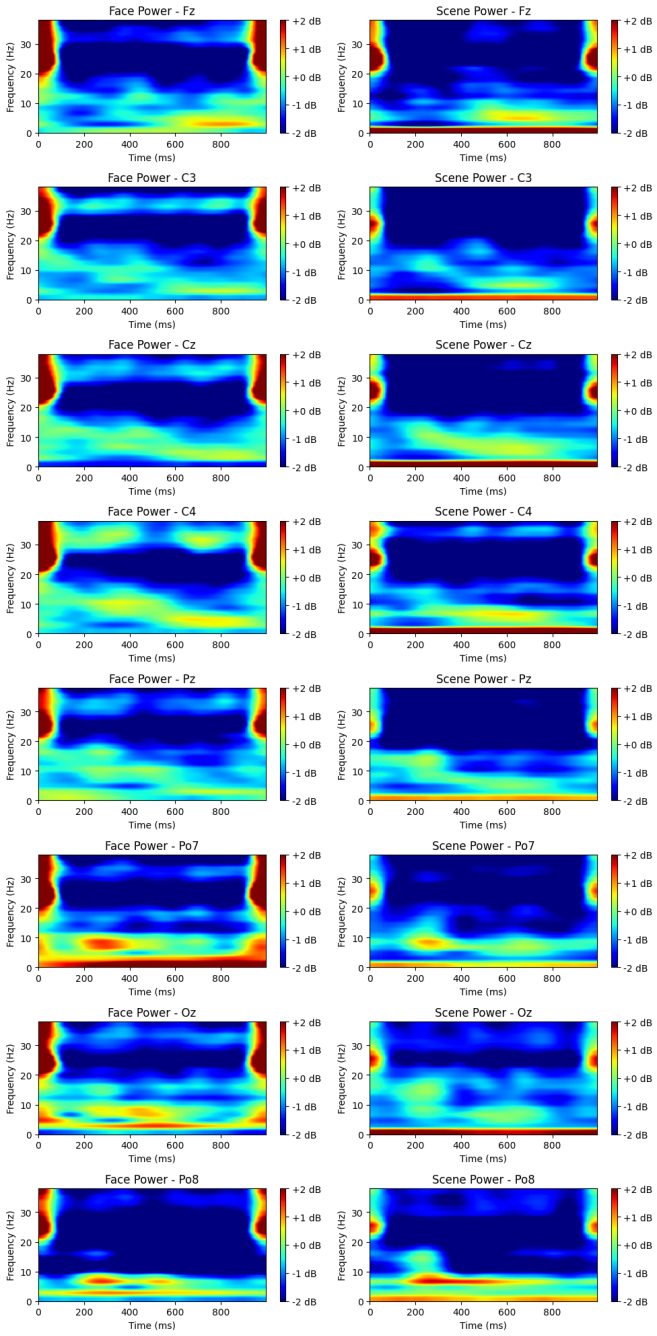
Figure 8 and Figure 9 show the Event-Related Potentials (ERPs) and time-frequency analyses of two additional human subjects, demonstrating the uniqueness of brain wave patterns among different individuals. They showcase the variability in neural responses to specific stimuli across different EEG channels. The ERP graphs display the voltage fluctuations over time, with distinctive waveforms for 'Scene', 'Face', and 'Base' conditions, highlighting how personal neural signatures can differ significantly between subjects. Similarly, the time-frequency plots reveal the power spectral density across various frequencies over time, offering a deeper view of the dynamic oscillatory processes of the brain during the experiment. These differences in ERPs and wavelet power spectra underscore the personalized nature of brain activity and the importance of individualized analysis in neuroscientific research.



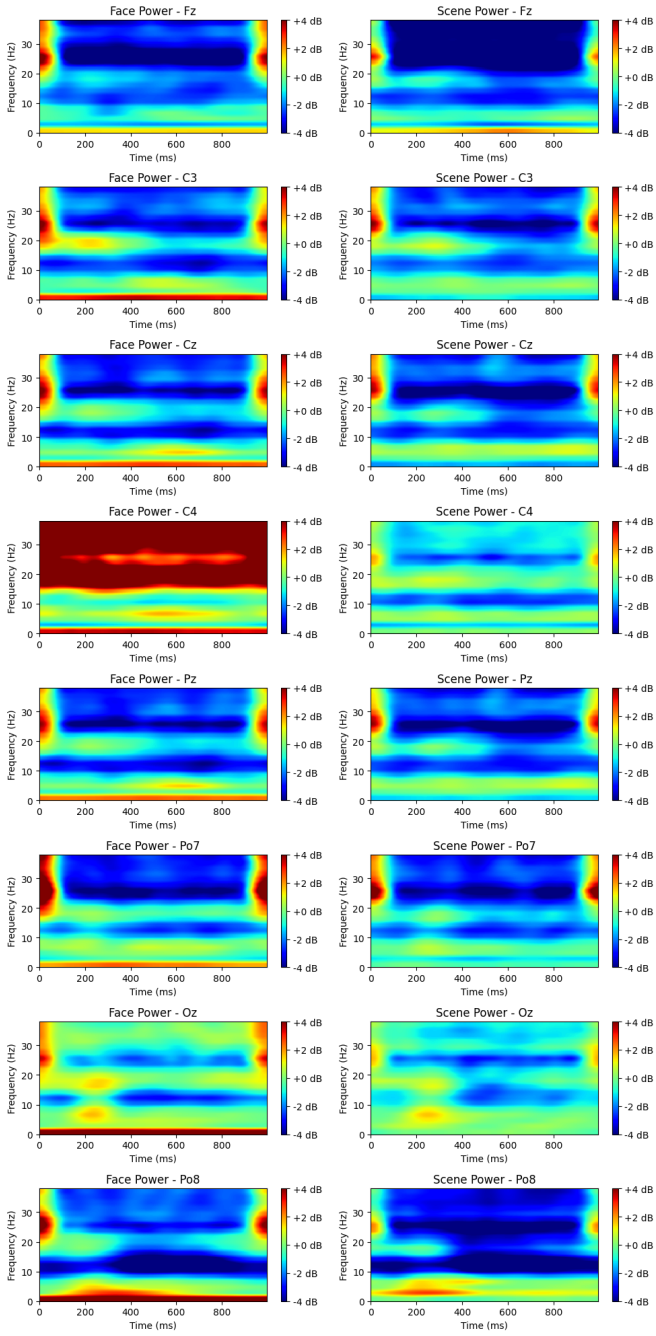
(a) ERP of the first subject

(b) ERP of the second subject

Fig. 8: Event-Related Potential (ERP)



(a) Time-Frequency Analysis of the first subject



(b) Time-Frequency Analysis of the second subject

Fig. 9: Time-Frequency Analysis