DECODING VISUAL STIMULI AND VISUAL IMAGERY INFORMATION FROM EEG SIGNALS UTILIZING MULTI-PERSPECTIVE 3D CNN BASED HIERARCHICAL DEEP-FUSION LEARNING NETWORK

By

Fatma Y Emanet

A Thesis Submitted to the Faculty
of Southeastern Louisiana University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Integrated Science and Technology

Southeastern Louisiana University
Hammond, Louisiana

May 2023

DECODING VISUAL STIMULI AND VISUAL IMAGERY INFORMATION FROM EEG SIGNALS UTILIZING MULTI-PERSPECTIVE 3D CNN BASED HIERARCHICAL DEEP-FUSION LEARNING NETWORK

By

Fatma Y Emanet

Approved:

Kazim Sekeroglu, Ph.D.
Assistant Professor of Computer Science
(Director of Thesis)

John Burris, Ph.D.
Associate Professor of Computer Science
(Committee Member)

Paulo Alexandre Regis, Ph.D.
Assistant Professor of Computer Science
(Committee Member)

David Gurney, Ph.D.
Assistant Professor of Mathematics
(Committee Member)

Daniel Mc Carthy, Ph.D.
Dean of the College of Science and Technology

Name: Fatma Y Emanet

Previous degrees: B.S., Bogazici University, 2019, (Computer Education and Educational Technology)

Date of Current Degree: May 13, 2023

Institution: Southeastern Louisiana University

Major Field: Integrated Science and Technology (Computer Science)

Major Professor: Dr Kazim Sekeroglu

Title of Study: DECODING VISUAL STIMULI AND VISUAL IMAGERY INFORMATION FROM EEG SIGNALS UTILIZING MULTI-PERSPECTIVE 3D CNN BASED HIERARCHICAL DEEP-FUSION LEARNING NETWORK

Pages in Study: 90

Candidate for Degree of Master of Science

The Brain-Computer Interface Systems (BCIs), a technology that aims to establish communication between the brain and machines. BCIs deal with brain activities and turn them into meaningful information, which can be used for developing brain-computer interfaces, diagnosing disorders, understanding human brain function, and augmenting cognition. This study aims to develop a model that utilizes brain decoding for the purpose of classifying not only visual stimuli but also visual imagery. Specifically, electroencephalographic (EEG) data collected from participants in an experimental setup designed for this study will be utilized for classification purposes. The main three objectives of this study are analyzing visual stimuli EEG signals and visual imagery EEG signals, exploring 2D spatiotemporal EEG image mappings for feature extraction and classification, and utilizing a 3D convolutional neural network (CNN) based multi-perspective hierarchical deep fusion model for the classification

of visually evoked signals as well as visual imagery signals represented by 3D EEG. In general, the experimental findings suggest that a fusion architecture exhibits superior performance compared to a model that operates independently.

Keywords: Hierarchical Deep Learning, Brain Computer Interface, Fusion Learning, Spatiotemporal Pattern Recognition, Multi-Perspective Learning

DEDICATION

I would like to dedicate this research to my husband and my baby boy.

ACKNOWLEDGMENTS

I would like to express my gratitude to the individuals who played an important role in the completion of this work. Firstly, I would like to thank Dr. Kazim Sekeroglu, my advisor, for his generosity in dedicating his time and effort to guide and assist me throughout the complexities of the master's program and the thesis process. In addition, I am grateful to my father, Prof.Dr. Ercan Oztemel, and my mother, Gulumser Oztemel, for their constant support in all aspects of my life. Lastly, I would like to extend my appreciation to my brother, Muhammed Esad Oztemel, for encouraging me to pursue this master's program.

The research reported in this thesis was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103424-21.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF GRAPHS

LIST OF CHARTS

CHAPTER I

INTRODUCTION

The brain-computer interface systems (BCIs) are one of the crucial technologies in recent years that aim to establish communication between the brain and machines. Besides the use of BCI systems in many scientific research areas, the main purpose of BCI systems is to enable people to develop applications where they can control various devices including computers, prosthetic limbs, robots, and even video games by using only power of human thought [1].

BCI is a system that deals with the brain activities of a living thing (human or animal) and turns these activities into meaningful information about the cognitive, perceptual, or motor processes associated with neural activity patterns. This process is also known as brain decoding. Meaningful knowledge obtained thanks to brain decoding can be used for studies such as developing brain-computer interfaces, diagnosing disorders, understanding human brain function, and even augmenting cognition [2].

In recent years, BCI technologies have started to show their presence in fields such as medicine, neuroscience, and gaming and are used for revolutionary innovations in these fields. Especially thanks to the BCI innovations made in the medical world, many people with disabilities and limited mobility have started to meet their needs without the need for any physical activity [3].

With many applications developed so far, this field frequently updates itself and is very open to developments. Therefore, it has the potential to understand the brain and its working principles, which is increasing day by day. This potential has attracted the attention of scientists and it has recently become a hot topic in the world of science and technology.

We can give examples of these applications; communication devices for people with disabilities [4], controlling devices in hazardous environments [5], enhancing cognitive performance [6], prosthetic limbs that can be controlled by the user's thoughts [7], and even brain-controlled video games [8].

In addition to many detailed and successful studies conducted in this area, developing reliable and robust decoding algorithms, and obtaining consistent neural activity patterns by brain decoding is still a challenge today due to some reasons related to the brain such as the complexity of the brain signals due to its nature, its dynamic structure, and being affected by environmental factors.

BCI technology measures and interprets brain activity using a range of approaches, including invasive, non-invasive, and semi-invasive methods. Invasive methods require implanting electrodes directly into the brain like electrocorticography (ECoG) and single-unit recordings, whereas non-invasive approaches record brain activity using external sensors such as EEG (Electroencephalography), fMRI, and others (Functional Magnetic Resonance Imaging). Semi-invasive methods include both invasive and non-invasive procedures [9].

Brain decoding can be used for visual stimuli classification. Visual stimuli classification refers to the process of identifying the category or features of a visual stimulus such as an image or video clip. It uses the response of the brain which is the patterns of neural activity that stimuli evoke in the brain for identifying the category of a visual stimulus [10].

Using various machine learning techniques, models with high performance can be created and successful results can be obtained to classify visual stimuli. These models can predict the category of a new, unseen visual stimulus based on the feature map which is most relevant for the cognitive task at hand by training the algorithm with known categories of visual stimuli, such as images of faces, letters, or simple shapes [11].

This study aims to develop a model that utilizes brain decoding for the purpose of classifying not only visual stimuli but also visual imagery. Specifically, electroencephalographic (EEG) data collected from participants in an experimental setup designed for this study will be utilized for classification purposes.

PROBLEM DEFINITION

Brain-computer interfaces are known for their potential to provide solutions to a wide range of issues in both scientific and everyday contexts. The primary objective of this project was to address a research-based problem related to the classification of EEG signals within the context of brain-computer interfaces. While this project was focused on addressing this specific issue, the insights and findings obtained through this work could be applied to a broader range of problems in various domains. In addition, the main three objectives of this study are as follows:

1. Analysis of visually evoked EEG signals and visual imagery EEG signals based on 2D spatiotemporal EEG image representation.
2. Exploring 3D EEG data consists of 2D spatiotemporal EEG image mappings (2D ST-EEG maps) for feature extraction and classification.
3. Utilizing a 3D convolutional neural network (CNN) based multi-perspective hierarchical deep fusion model for the classification of visually evoked signals as well as visual imagery signals represented by 3D EEG.

MOTIVATION

A brain-computer interface (BCI) is a technological system that enables the transfer of brain signals or patterns from a user to an interactive application. The use of BCI allows a user to control the interactive application by only relying on their imagination, for instance, by commanding the cursor to move left or right across the screen. Given its potential, BCI plays a crucial role in facilitating the control of assistive technologies, such as text input systems, wheelchairs, and rehabilitation devices for stroke patients, especially among those with motor impairments.

BCI is a newly emerging area of research that has gained considerable attention in academia. While some valuable studies exist in this field, they are limited in scope, highlighting the need for further exploration and study.

As a result, with guidance from my thesis advisor, I chose to pursue my master's studies at BCI. Specifically, I focused on brain decoding, visual stimuli classification, and visual imagery classification as the primary areas of concentration in a research project led by my advisor. The proposed method in this study can be used for recognizing EEG signals for BCIs as well as for the diagnosis of neurological disorders and diseases.

DATASET

Electroencephalography (EEG) technology is used to obtain brain activity patterns in this study. Rather than utilizing an existing dataset, EEG data were collected from volunteer participants using an EEG tool, resulting in the creation of a new dataset for analysis purposes.

The name of this device is “Enobio 32”. This device is produced by the “NEUROELECTRICS” Company [12]. This company is a medical company in the field of non-invasive brain stimulation technology for personalized neuromodulation. Our preferred EEG tool in our experimental setup to collect data is shown in Figure 1.

Figure 1: Enobio 32 EEG Tool

The Enobio 32 system is a medical device that can be used in clinical procedures as well as clinical research. There are 32 electrodes on the device, which means that it can collect brain signals from 32 different channels. Two setup options, namely dry and gel electrodes, are available. The dry electrode option was chosen for our study to save time.

Equipped with a rechargeable battery, the Enobio 32 system is capable of operating for up to 5.5 hours on a single charge. It supports wireless data transmission within a 10-meter range from the computer running the controller software NIC2. Furthermore, the device is equipped with an SD card recording feature, known as holter mode, which enables mobile experiments lasting up to 19 hours without requiring a computer.

The wireless functionality of the Enobio 32 system was utilized in the experiment, and the NIC software that was included in the device's package was employed for the data recording phase. This software not only allows real-time visualization of brain signals during signal acquisition but also offers a plethora of analysis tools, including scalp maps, spectrum analysis, and cortical maps. Consequently, the NIC software provides an opportunity to monitor these analysis tools in real-time during the experiment. Examples of these analyses can be found in Figures 2 and 3.

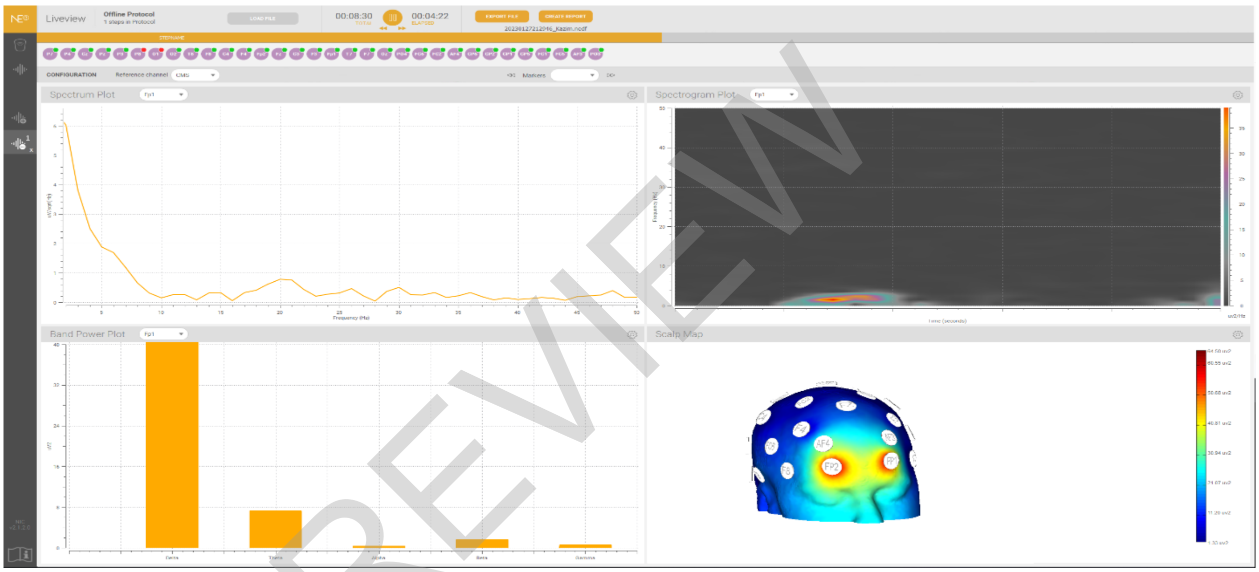

Figure 2: NIC software analysis tools

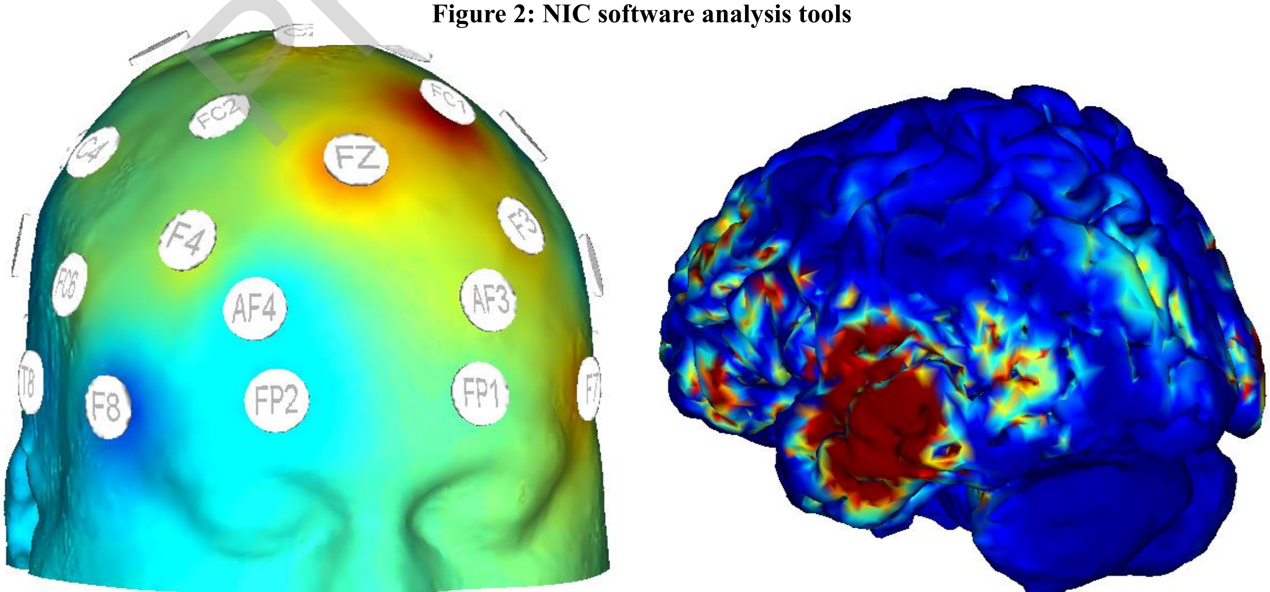

Figure 3: NIC software analysis tools

The sampling rate is fixed at 500. A 10-10 international electrode placement system is used in the device. Figure 4 shows the international 10-10 system [13].

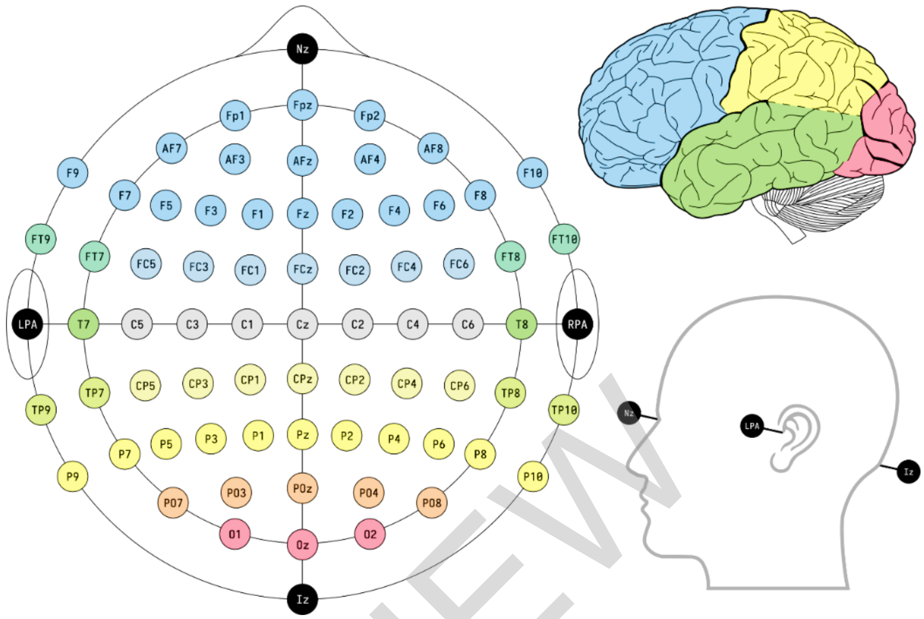

Figure 4: International 10-10 system of electrode placement

The data is collected from 16 different adult volunteer participants with normal or corrected normal vision. The participants were informed about the experimental protocol and signed the informed consent form before the experiment. It was collected in 4 setups on different days. A total of over 10.000.000 brain signals were collected during this process.

Some preliminary experiments were carried out before preparing the final experimental setup. First, the participants were presented with a letter on the screen for a duration of 5 to 10 seconds. Following this, they were instructed to visualize the letter while looking at a black screen. The aim of this exercise was to determine the feasibility of decoding visual and imagined visual information from EEG signals, using the proposed technique. In our initial data collection, all 26 letters were presented to each participant, and they were asked to imagine each one. Upon analysis, it was observed that every participant exhibited a consistent EEG pattern, represented by 2D ST-EEG maps when looking at the same stimulus. However, it was noted that the 2D EEG mapping differed across different participants for the same visual or imagined stimulus. These results prompted us to conduct a second round of data collection, during which we presented the same letter multiple times in a randomized order. Our aim was to investigate whether there was a consistent pattern in the subject's brain activity when they viewed the same stimulus at different times. This would allow us to develop a personalized model capable of accurately predicting and decoding EEG signals for that specific individual.

Following the second iteration, it was noted that a greater number of classes required more complex models necessitating a larger dataset. Consequently, to lay the groundwork for future investigations, it was decided to commence with a smaller number of classes as preliminary work. As a result, a third iteration was conducted, this time with only four classes, consisting of the letters "A", "B", "C", and "Rest". These letters were chosen due to their geometric dissimilarities. Based on the outcomes of this iteration, the final experimental setup was developed.

In our final experimental setup, data was collected from seven volunteer participants to train the proposed model, resulting in a total of 5,379,500 signals collected. Specifically, 418,500 signals were collected in the first section and 350,000 signals were collected in the second section for each individual participant.

Experimental Setup

The participants were seated in a quiet, well-lit room to minimize any potential distractions. They were comfortably seated in a chair positioned 50 cm away from an LCD computer screen. Prior to starting the experiment, the participants were instructed to remain seated in the chair for the duration of the experiment, during which an EEG cap was placed on their heads to record their brain waves.

The experiment consisted of two sections, each conducted on the same day. During each section, the EEG recorder recorded the participants' brain waves while a stimulus was randomly displayed via a slideshow. To maintain consistency between the two sections, the EEG cap remained on the participants' heads during the interim period. The stimuli presented in both sections were the letters A, B, and C randomly displayed.

In the first experiment section, participants were instructed to first fix their gaze on the computer screen with their eyes open for 30 seconds, followed by 30 seconds of eyes-closed rest. Next, the letters A, B, and C were randomly displayed on the screen for 10 seconds each. Before the presentation of the next letter, participants were shown a black screen for 1.5 seconds to allow their brain signals to return to a resting level and minimize the influence of the previous letter. This process was repeated for each letter 20 times in random order.

In the second section, participants were presented with a blank white screen for the entire duration. An auditory cue instructed participants to imagine a specific letter, which they were asked to maintain in their minds for 10 seconds. After this, participants were instructed to imagine another randomly chosen letter for 10 seconds. This process was repeated 20 times for each of the letter A, B, and C, resulting in a total of 20 imagined instances of each letter. The process of each experiment section is illustrated in Figure 5 and 6. The next page displays a collage of photographs taken during data collection in Figure 7 and 8.

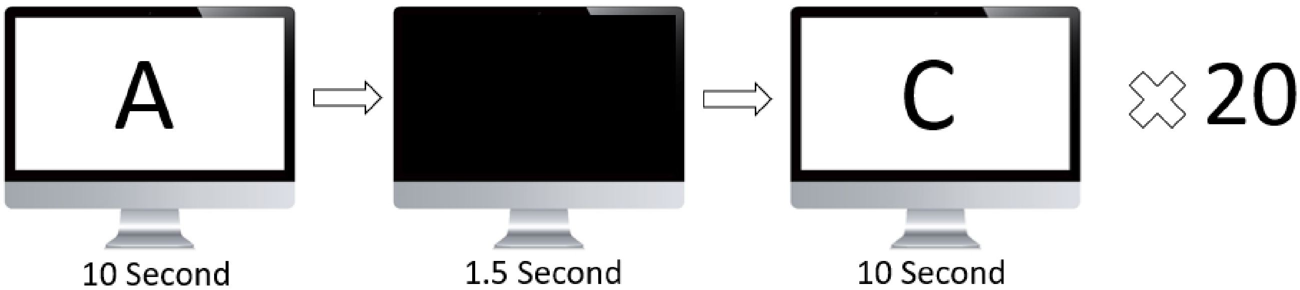

Figure 5: The First Experiment Section-Visual Stimuli Phase

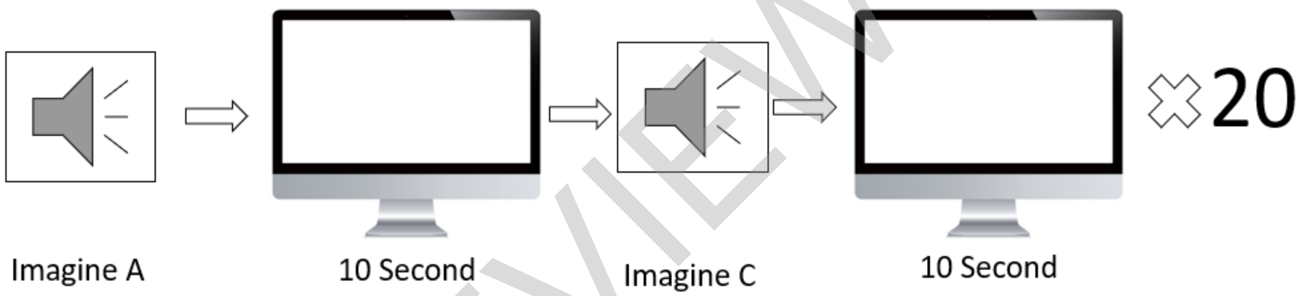

Figure 6: The Second Experiment Section-Visual Imagery Phase

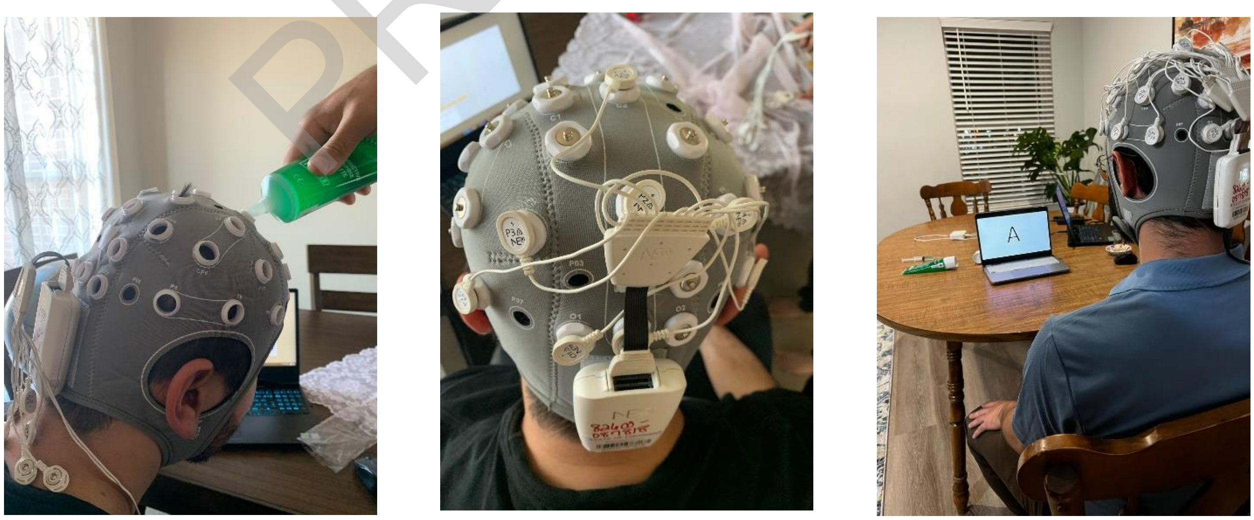

Figure 7: Data Collection I

Figure 8: Data Collection II