

## **A Design of Systems**

in Image & Auditory Reconstruction From EEG Brain Signals



Graduate Project 1 (3183) - Karpjoo Jeong

Team 7, EEmage – Advisor : Eun Yi Kim

Computer Science and Engineering 201811218 Hyun Woo Lee

Computer Science and Engineering 201811182 Wonjun Park

Computer Science and Engineering 201911193 Esan Woo

## Abstract

이 보고서는 EEG (Electroencephalography) Brain Signals로부터 인간의 시각 이미지와 청각 소리를 재현하는 연구의 시스템 설계서입니다. EEG data는 Brain-Computer Interface (BCI) 분야에서 각광받고 있는 비용 효율적인 데이터 중 하나로, 이를 통해 시각을 재현할 수 있다는 것은 기정사실화 되어 가고 있으며, 청각에 대해서도 그 시도가 이어지고 있습니다. 이 보고서에서는 사용하는 프레임워크와 데이터셋을 소개합니다.

## I. Introduction

본 연구의 시발점은 현재 “인간의 시각 체계가 형성되는 데에는 Supervision이 필요하지 않다”는 가설을 토대로 합니다. 이 가설을 통해 시각 이미지 재현 프레임워크 중 인간의 시각 체계에 해당하는 인코딩 방법에 Supervised Learning을 사용하는 것이 실제 인간과 맞지 않다고 생각하여, 궁극적으로 이를 탈피하는 새로운 프레임워크를 제안하고자 합니다.

그 과정에서 A Universal EEG Encoder에 대한 생각을 하게 되었고, 그 하위 목표로서 2023년 12월에 마감되는 ICASSP 2024: A EEG Auditory Challenge [1] 를 먼저 준비하게 되었습니다. 해당 challenge는 두 가지 task가 주어집니다. 첫 번째는 Match-Mismatch로 하나의 소리 자극을 다섯 개의 부분으로 나누어 그 중 정답을 고르는 것입니다. 인공지능 모델은 주어진 EEG 데이터를 가지고, 해당 데이터가 어떤 소리에 해당하는 지 골라야 하는 classification 문제입니다. 두 번째는 인공지능 모델이 주어진 EEG 데이터를 가지고 Mel-spectrogram을 만드는 Regression 문제에 해당합니다.

이 challenge를 통해, 청각과 EEG data 간의 상관도를 분석, 파악할 수 있을 것이라 기대합니다. 이는 범용 인코더를 만들기 위한 초석이 될 것입니다.

## II. System Overview

본 연구에서는 해당 챌린지를 두 가지 방향에서 접근하고 있습니다. 먼저, 각 EEG 채널에서 나오는 데이터의 의미를 파악합니다. 다른 방법은 시각 재현 프레임워크를 적용하는 방법입니다. 각 방법들에 대해서는 다음의 세부주제에서 서술합니다.

### A. Analyze EEG Domain Data

EEG를 통한 뇌파 분석에 대해 이해한다면 모델을 설계하는데 큰 도움이 될 것입니다. 실제로 fMRI를 이용한 방법은 비용 측면에서 비효율적이지만 그 효과를 여러 연구에서 입증했습니다. 해당 연구들을 토대로 추측한 청각 관련 Channel들은 다음 Figure에 노란색과 파란색으로 표시하였습니다.

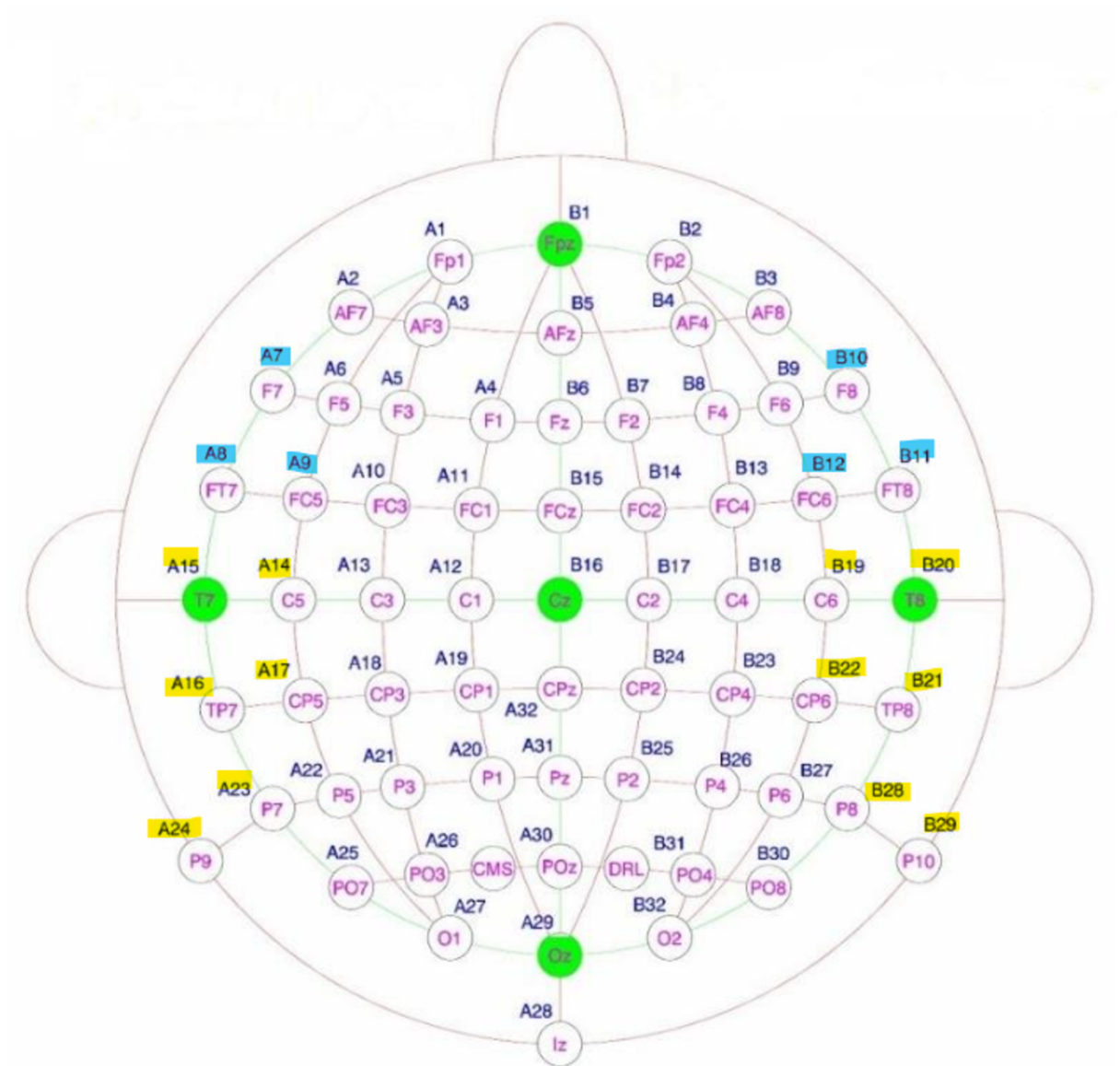


Figure 1. EEG Channels for the 64 Environment

또한 각 Channel별로 Variance와 Pearson Correlation 등의 수치를 찍어보며 의미있는 Channel을 찾아내는 데에 초점을 맞추고 있습니다. 찍어낸 Variance는 다음과 같습니다.

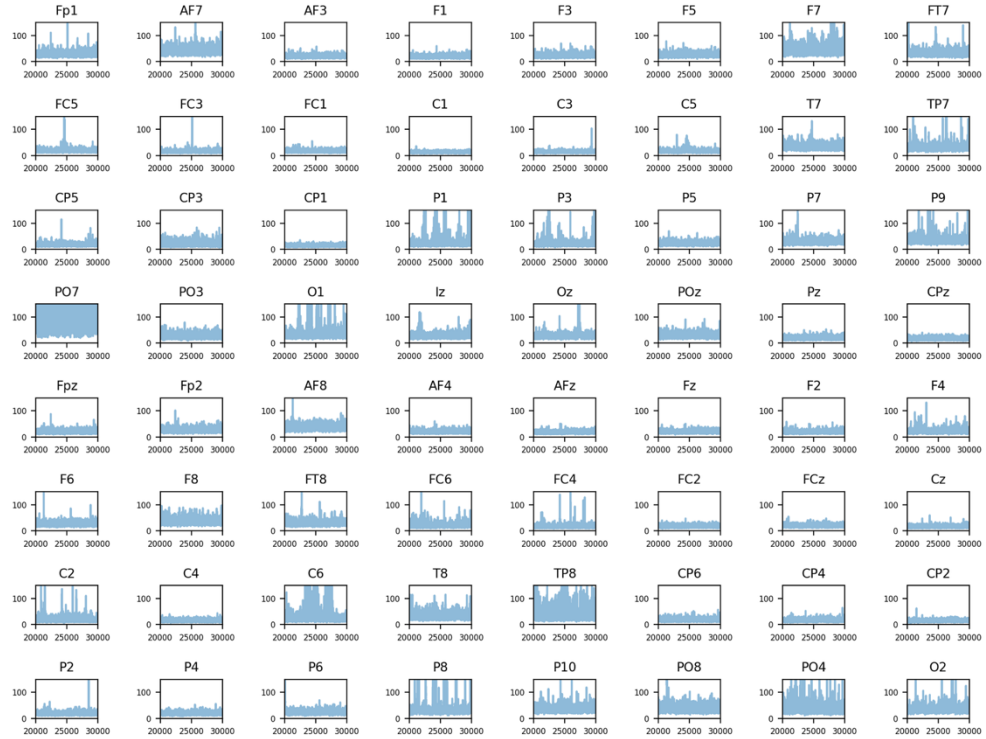
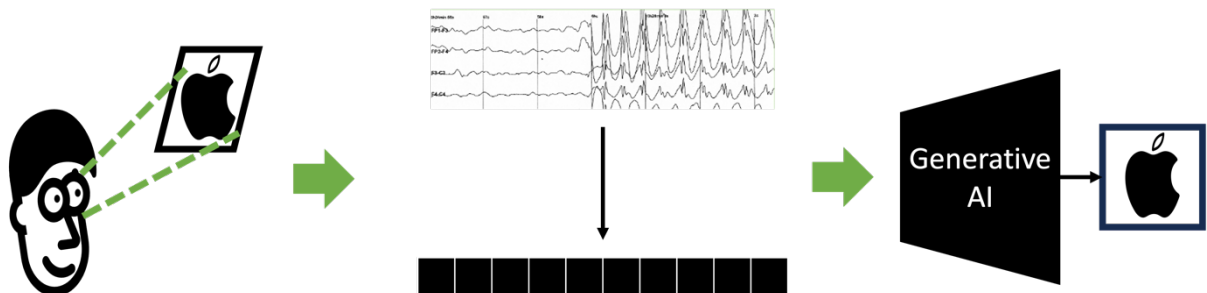


Figure 2. Variances for Each EEG Channel

이외에도 다양한 데이터 분석 방법을 통해 데이터를 분석할 예정입니다.

## B. Applying the Framework of The Human Vision Reconstruction

이미지와 소리는 확실히 다른 도메인에 속합니다. 하지만 [2], [3] 와 같이, 이 두 도메인을 함께 분석하려는 시도는 계속해서 이루어지고 있습니다. 본 연구 또한 해당 연구들과 같은 “이미지와 소리 데이터는 함께 학습될 수 있다” 는 가정을 내포합니다. 따라서 비교적 활발하게 이루어진 Vision Reconstruction 방법론을 Auditory Reconstruction에 활용해보고자 합니다. 사용하는 프레임워크는 다음과 같습니다.



이때 Feature Vector를 뽑아내는 데 사용되는 것을 Encoder라 합니다. 현재는 제안된 여러 이미지 Encoder 중 EEG-ChannelNet [4] 을 활용해보려고 준비하고 있습니다.

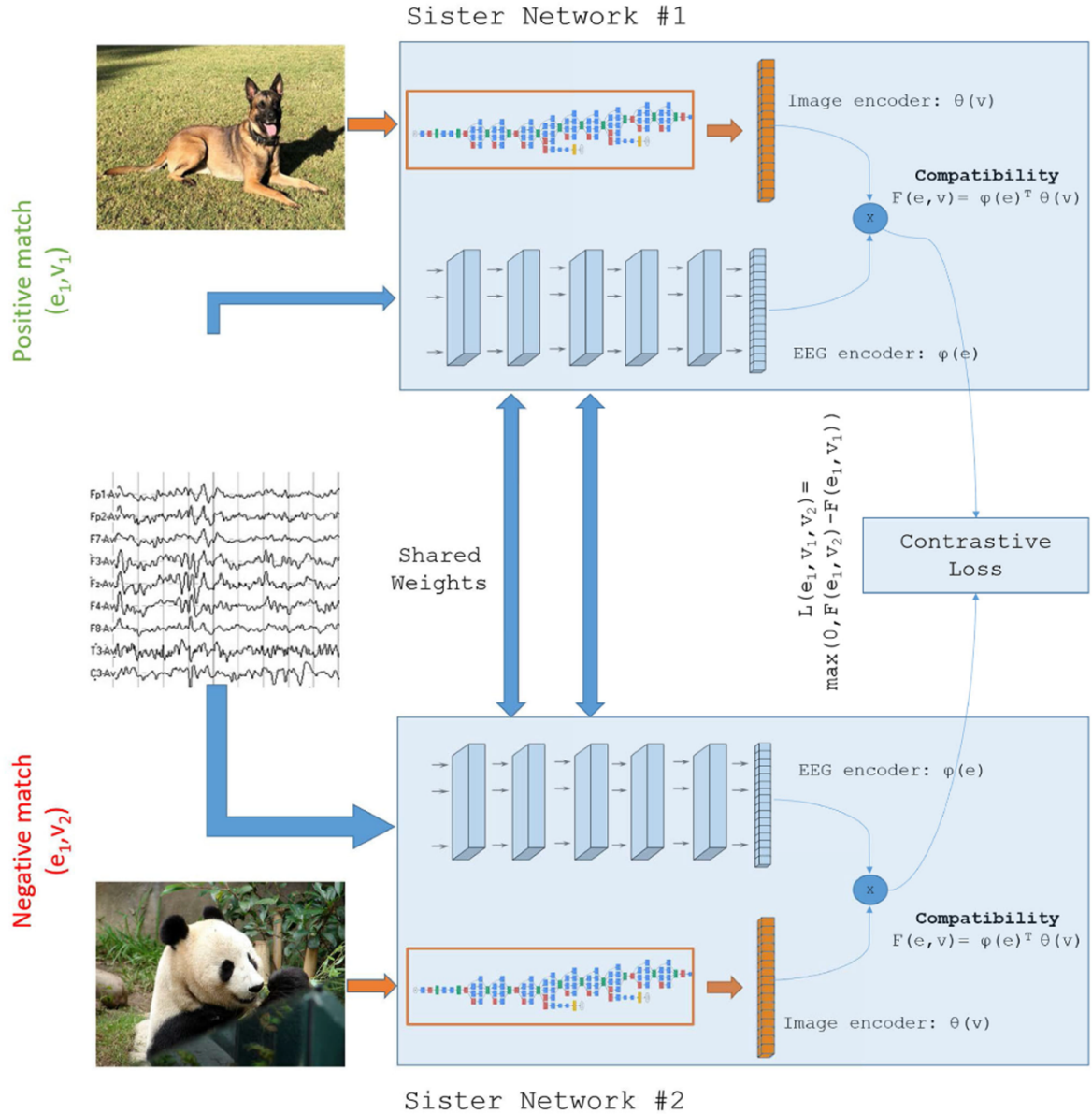


Figure 3. Multimodal Learning Framework in Image Reconstruction

(4) 는 Figure 3 의 프레임워크에서 EEG Encoder에 EEG-ChannelNet을 사용했습니다. EEG 와 매칭되는 이미지와 그렇지 않은 이미지를 동시에 학습시키는 방법을 사용했습니다. Loss function은 Triplet Loss를 이용하여 EEG와 매칭되는 이미지와 그렇지 않은 이미지간의 차이를 크게 만들었습니다.

### III. Dataset

(1) 에서 사용되는 데이터셋 SparrKULee [5]는 85명의 참가자들로부터 약 90-150분의 Auditory Stimuli를 주어 EEG Data를 수집한 것입니다. 모든 참가자는 AudioBook1이라는 공통된 Stimuli를 들었으며, 그 외의 라벨(Sessions, various from Ses-Shortstories01 to Ses-Varyingstories10)에 대해서는 모두 다른 Stimuli를 듣도록 했습니다. 각 라벨은 최소 두 명부터 26명까지 참가자 수가 다양합니다.

EEG Data는 제공된 전처리 코드는 최종적으로 64Hz로 Downsample 하여 사용합니다.

하지만 II.A. 에서 다른 방식으로 전처리를 하는 것이 더 유의미한 데이터를 만들어낼 수 있다는 것을 밝힌다면 다른 방법 또한 고려할 예정입니다.

#### IV. Conclusion

본 프로젝트는 궁극적으로 인간의 시각과 청각을 재현할 수 있는 Universal EEG Encoder의 제안을 목표로 합니다. 그 과정에서 EEG, Image, Auditory 데이터의 상관관계를 파악하고자 ICASSP 2024: An Auditory Challenge를 준비합니다. 이는 본 연구에서 사용할 수 있는 데이터의 범주를 늘려, 제안하고자 하는 인코더의 성능을 높일 수 있는 계기로 작용할 것이라 기대합니다.

#### Reference

다음의 목록은 MLA 포맷을 따랐습니다.

- [1] "Auditory EEG Challenge 2024." *ICASSP 2024 SP Grand Challenges*, <https://exporl.github.io/auditory-eeg-challenge-2024/>.
- [2] Likhoshesterov, Valerii, et al. "Polyvit: Co-training vision transformers on images, videos and audio." *arXiv preprint arXiv:2111.12993* (2021).
- [3] Gong, Yuan, et al. "Uavm: Towards unifying audio and visual models." *IEEE Signal Processing Letters* 29 (2022): 2437-2441.
- [4] Palazzo, Simone, et al. "Decoding brain representations by multimodal learning of neural activity and visual features." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.11 (2020): 3833-3849.
- [5] Accou, Bernd, et al. "SparrKULee: A Speech-evoked Auditory Response Repository of the KU Leuven, containing EEG of 85 participants." *bioRxiv* (2023): 2023-07.