Neural Encoding for Dynamic Natural Vision

Project Report Submitted
to
MANIPAL ACADEMY OF HIGHER EDUCATION

For Partial Fulfillment of the Requirement for the
Award of the Degree
Of
Bachelor of Technology
in
Computer and Communication Engineering
By

KUNAL YADAV
Reg. No.– 210953130
NAMYA PATIYAL
Reg. No.– 210953062

Under the guidance of

Ms Chethana Pujari
Assistant Professor-Senior Scale
Department of I&CT
Manipal Institute of Technology
Manipal, Karnataka, India

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
A Constituent Unit of MAHE, Manipal

Table of Contents

INTRODUCTION

In the ever-evolving landscape of neuroimaging and brain-computer interfaces, the intersection of artificial intelligence and neuroscience has paved the way for groundbreaking research aimed at deciphering the intricate language of the human brain. Our project, resides at the confluence of neuroscience, machine learning, and neuroimaging, seeking to unravel the complex patterns within intracranial electroencephalogram (iEEG) data to predict visual stimuli perceived by the human brain.

The fundamental premise of our endeavor lies in the hypothesis that the electrical activity of the brain, as captured through iEEG recordings, encapsulates discernible patterns that mirror the visual stimuli presented to the subject. This endeavor not only delves into the nuances of neural encoding but also capitalizes on the capabilities of advanced machine learning models, particularly Convolutional Neural Networks (CNNs), to bridge the gap between raw brain signals and the corresponding visual experiences.

The project spans multiple domains, reflecting a multidisciplinary approach essential for tackling the complexity inherent in decoding neural responses to external stimuli. We traverse the realms of neuroscience, exploring the intricacies of how the brain processes and responds to visual information. Simultaneously, we navigate the landscape of machine learning, employing sophisticated algorithms to discern meaningful patterns from the vast and intricate iEEG datasets.

The challenges encountered in this project are manifold, ranging from the sheer size of the data, which necessitates innovative preprocessing strategies, to the inherent complexity of mapping neural patterns to visual stimuli. Moreover, the study acknowledges the memory constraints inherent in handling large iEEG segments and visual stimuli images during model training.

Understanding the intricate relationship between neural activity and external stimuli is a fundamental pursuit in neuroscience. This research explores this connection by harnessing the power of electro-corticography (ECoG) data and visual stimuli processing. The primary objective is to develop a predictive model capable of discerning visual stimuli from intracranial electroencephalography (iEEG) recordings, hypothesising that discernible patterns in the brain's electrical activity correlate with the visual experiences of the subject.

LITERATURE SURVEY

1. Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies

Authors: S. Nishimoto et al. (2011)

Problem Statement: Decoding brain responses to dynamic visual stimuli, particularly natural movies, is challenging. This study aims to understand and reconstruct visual experiences from brain activity evoked by natural movies.

Methods: Functional Magnetic Resonance Imaging (fMRI) was used to capture brain activity while viewing natural movies. A two-stage model translated movies into fMRI data and reconstructed visual stimuli from this data.

Advantages: Tackling dynamic stimuli, a two-stage model for nuanced understanding, and the ability to generalize to unseen data.

Disadvantages: Computational complexity and a focus primarily on the visual cortex.

Conclusion: The study contributes significantly to understanding how the brain processes dynamic visual stimuli, providing insights with potential applications in various domains.

2. Identifying natural images from human brain activity

Authors: K. N. Kay et al. (2008)

Problem Statement: Identifying specific natural images based on brain activity poses a complex challenge. This study investigates the possibility of predicting the image a person is looking at based solely on their brain activity.

Methods: Functional Magnetic Resonance Imaging (fMRI) recorded brain activity while participants viewed natural images. A computational model correlated specific brain activity patterns with individual images.

Advantages: High-resolution mapping, predictive modeling, and a focus on natural images.

Disadvantages: Limitations of fMRI, potential model specificity, and the challenge of capturing rapid neural responses.

Conclusion: The study demonstrates the feasibility of predicting viewed images from brain activity, with potential applications in brain-computer interfaces and diagnostics.

3. Reactivation of visual-evoked activity in human cortical networks

Authors: M. Chelaru et al. (2016)

Problem Statement: Understanding how the human brain reactivates visual-evoked activity is crucial for comprehending neural pathways. The study investigates reactivation patterns within human cortical networks.

Methods: fMRI recorded brain activity during exposure to specific visual stimuli, focusing on understanding reactivation patterns upon repeated exposure.

Advantages: Deep insights into neural pathways, focus on reactivation, and implications for memory and recognition mechanisms.

Disadvantages: Complexity of neural networks and inter-individual variability.

Conclusion: The study contributes to understanding how the brain processes, remembers, and reacts to visual stimuli, with implications for neuroscience and education.

4. Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data

Authors: T. Grootswagers et al. (2016)

Problem Statement: Decoding dynamic brain patterns from time series neuroimaging data is challenging. This tutorial provides a comprehensive guide on using Multivariate Pattern Analysis (MVPA).

Methods: The tutorial covers data preprocessing, model training, validation, and interpretation using MVPA, with practical examples.

Advantages: Comprehensive guide, practical examples, and insights into decoding dynamic brain patterns.

Disadvantages: Tutorial nature and potential limitations in addressing advanced MVPA techniques.

Conclusion: The tutorial demystifies MVPA for decoding brain patterns, offering practical insights for researchers working with time series neuroimaging data.

5. Decoding semantics from dynamic brain activation patterns: From trials to task in EEG/MEG source space

Authors: F. Magnabosco & O. Hauk (2023)

Problem Statement: Decoding semantic information from EEG/MEG data is a complex task. The study bridges the gap between individual trials and overall task performance in the EEG/MEG source space.

Methods: Advanced signal processing techniques were applied to analyze EEG/MEG data, focusing on source space and multivariate decoding.

Advantages: High spatial resolution, broad application, and potential for various research scenarios.

Disadvantages: Complexity, requiring a deep understanding of neuroimaging and signal processing, and potential data requirements.

Conclusion: The study provides a promising avenue for decoding detailed semantic information from brain activity, extending the applications of EEG/MEG data.

METHODOLOGY

Data Preprocessing:

The initial image data extraction resulted in images of shape 1030×1030×36, where 36 denotes the temporal dimension, representing time. To facilitate model training, these images were resized to 224×224 pixels. The corresponding temporal information, represented by durations such as 0.01667 and 0.03333 seconds, was retained.

Dataset used: https://openneuro.org/datasets/ds004194/versions/1.0.1

Temporal Segmentation:

An average stimulus duration of approximately 0.025 seconds was observed. To align with a desired segment duration of 1 second, we chose a 2-second window for each segment, capturing 1 second before and after the stimulus onset. This approach resulted in grouping approximately 40 stimuli together.

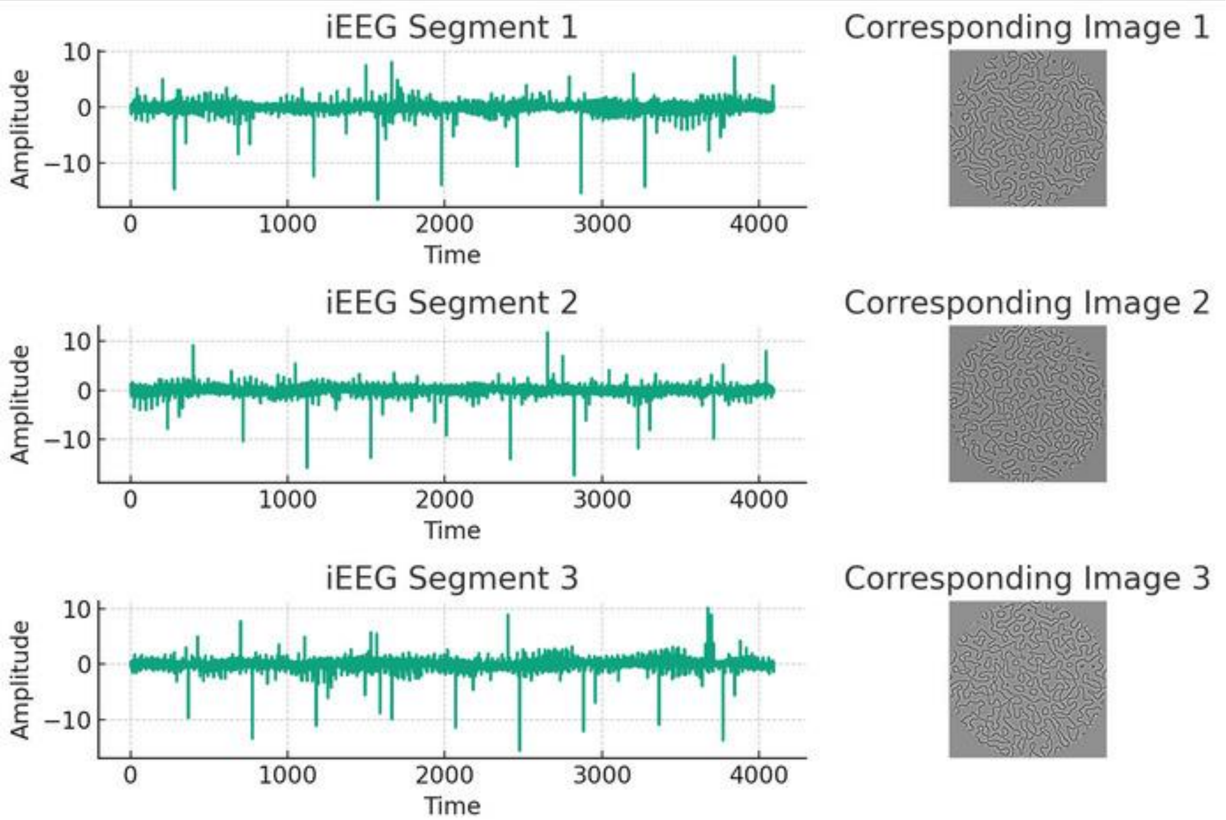

Fig. iEEG Segments of corresponding images

Data Pairing:

For each iEEG segment, we associated the visual stimulus presented during its onset. The visual stimuli, available as images in the stimulus data field, were resized to 224×224 pixels and normalized. As the number of stimuli images slightly exceeded the number of iEEG segments, we matched the first 36 images with the 36 iEEG segments, assuming a one-to-one correspondence.

Data Representation:

The paired iEEG segments and visual stimuli were transformed into a 224×224×36 array, serving as input and target data for model training.

Dataset Structure:

The dataset, located in the 'ieeg' subdirectory, contains various files for different tasks (prf, spatialpattern, temporalpattern) and runs. File types include .eeg, .vhdr, .vmrk, .json, and .tsv, providing information about iEEG recordings, metadata, channels, and event markers.

Task-Specific Files:

<u>Spatial Pattern Task</u>

- sub-p01_ses-umcuiemu01_task-spatialpattern_acq-clinical_run-01.mat
- sub-p01_ses-umcuiemu01_task-spatialpattern_acq-clinical_run-02.mat
- sub-p01_ses-umcuiemu01_task-spatialpattern_acq-clinical_run-03.mat
- sub-p01_ses-umcuiemu01_task-spatialpattern_acq-clinical_run-04.mat

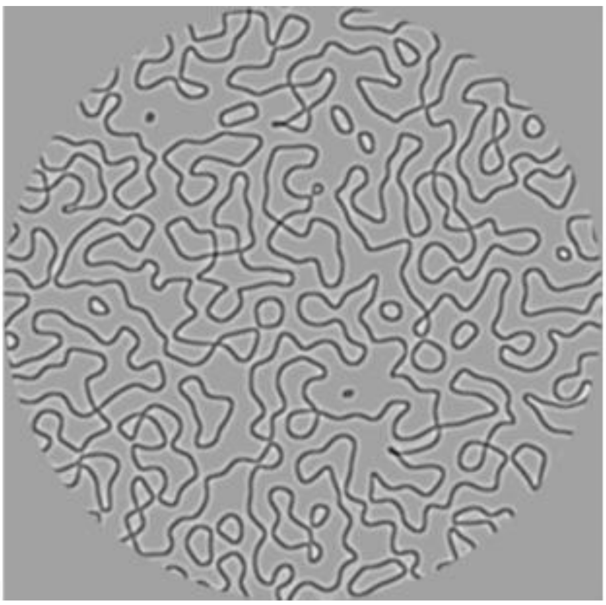

Fig. Task-Spatial Pattern – Image 1

Temporal Pattern Task

- sub-p01_ses-umcuiemu01_task-temporalpattern_acq-clinical_run-01.mat
- sub-p01_ses-umcuiemu01_task-temporalpattern_acq-clinical_run-02.mat
- sub-p01_ses-umcuiemu01_task-temporalpattern_acq-clinical_run-03.mat
- sub-p01_ses-umcuiemu01_task-temporalpattern_acq-clinical_run-04.mat

Metadata in .mat Files

The loaded .mat file for the temporalpattern task contains keys like params, quitProg, stimulus, ans, time0, timeFromT0, response, timing, pc, rc, pth, fname. These keys hold information about task parameters, program termination, stimuli, task outputs, timestamps, participant responses, and file details.

Stimulus Data in .mat Files

The stimulus key in the .mat file is a structured numpy array with fields like cmap, srcRect, dstRect, display, categories, images, im_cell, catindex, and duration. These fields provide details about color maps, stimulus positioning, display settings, stimulus categories, images, and durations.

iEEG Data Details

The iEEG data was recorded at a sampling frequency of 2048 Hz, crucial for processing. The data shape is 144,448 samples × 64 channels, with a recording duration calculated as the number of samples divided by the sampling frequency.

Temporal Alignment

Onset times of visual stimuli have been converted to sample indices in the iEEG data, enabling precise temporal alignment. The first few onsets, presented both in seconds and sample indices, exemplify this alignment:

3.0 seconds corresponds to sample index 6144

4.68 seconds corresponds to sample index 9591

6.32 seconds corresponds to sample index 12936

8.0 seconds corresponds to sample index 16384

9.67 seconds corresponds to sample index 19797

RESULTS AND DISCUSSION

The project aimed at reconstructing images using convolutional neural networks (CNN) and a long short-term memory (LSTM) networks,. The closest reproducible result achieved a similarity score of 0.307 on a scale of -1 to 1, indicating a moderately successful reconstruction.

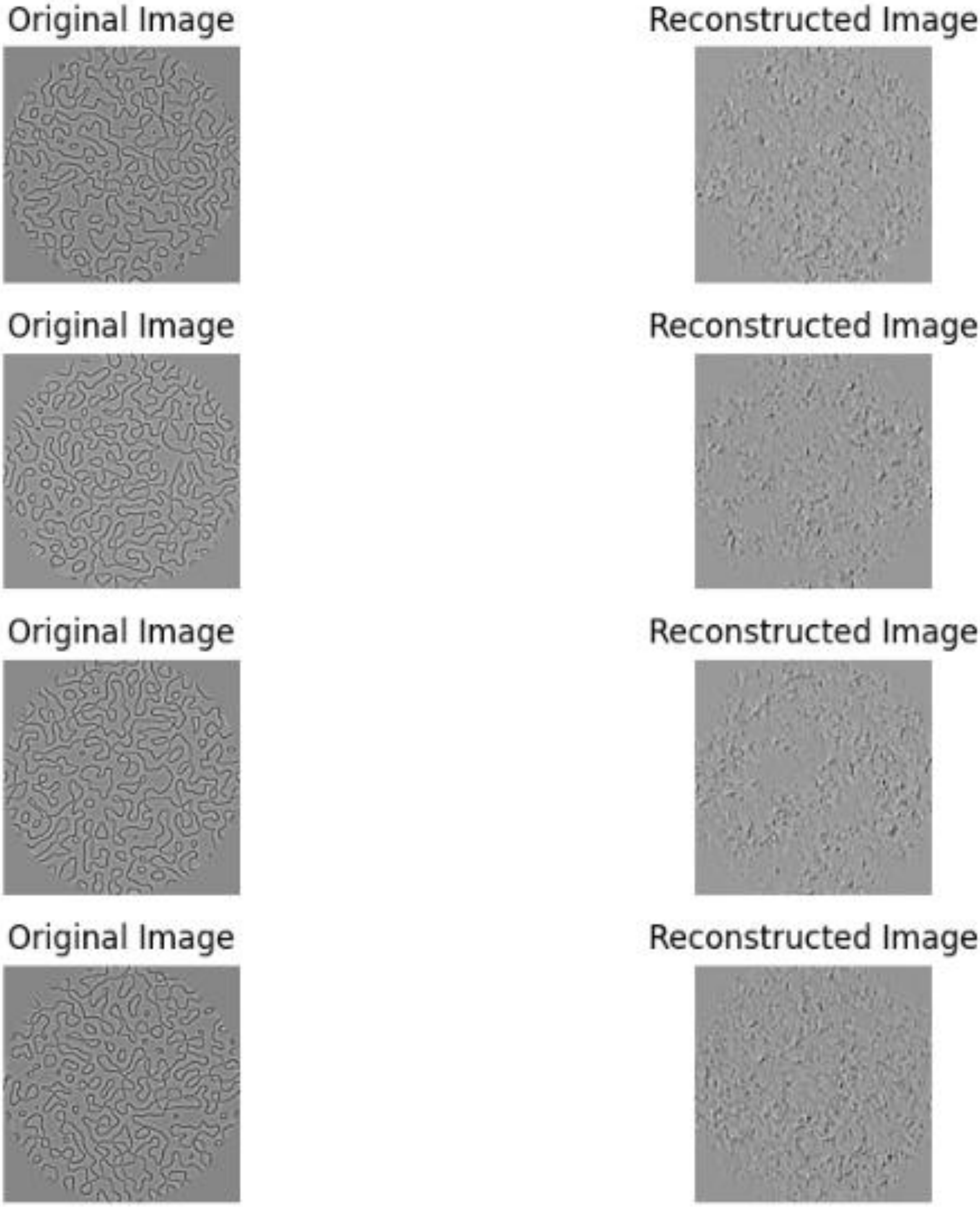

Fig. Original Image and Corresponding Reconstructed Image

In addition to the visual assessment, Grey Level Co-occurrence Matrix (GLCM) properties were calculated for both the original and reconstructed halves of the image. The results are as follows:

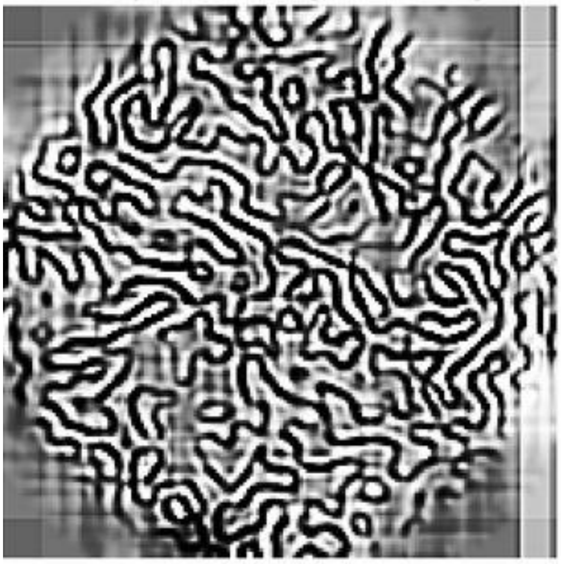

Fig. The original image

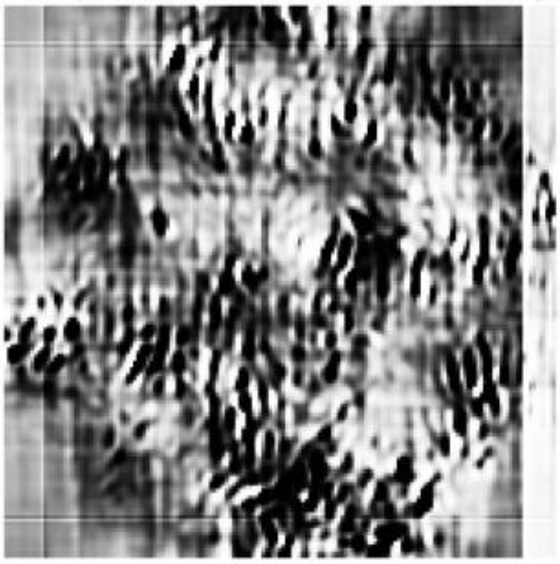

Fig. The reconstructed image

Original half:

Contrast: 14751.38
Dissimilarity: 87.07
Homogeneity: 0.156
Energy: 0.115
Correlation: 0.126
ASM (Angular Second Moment): 0.013

Reconstructed half:

Contrast: 8099.79
Dissimilarity: 61.77
Homogeneity: 0.126
Energy: 0.093
Correlation: 0.454
ASM (Angular Second Moment): 0.009

These GLCM features provide quantitative insights into the texture and spatial relationships within the images. While some features show differences between the original and reconstructed halves, further analysis is needed to interpret the implications of these variations.

Despite achieving a moderately successful reconstruction based on visual similarity, the GLCM analysis reveals differences in texture properties between the original and reconstructed halves. The decreased contrast and dissimilarity in the reconstructed image suggest a potential loss of detailed information compared to the original. The increase in correlation indicates a stronger linear relationship between pixel intensities in the reconstructed half.

However, it's important to note that GLCM features might not capture all aspects of visual similarity, and further investigation is needed to understand the perceptual implications of these differences. It would be beneficial to conduct a human perceptual study to assess the subjective quality of the reconstructed images.

CONCLUSION

This project undertook the formidable task of predicting visual stimuli from intracranial electroencephalography (iEEG) data using a Convolutional Neural Network (CNN). The results, both in terms of visual similarity metrics and Grey Level Co-occurrence Matrix (GLCM) analysis, provide valuable insights into the intricate relationship between brain activity and visual perception.

The achieved similarity score of 0.307 on a scale of -1 to 1 signifies a promising step forward in decoding visual stimuli from iEEG data. The GLCM analysis, revealing variations in texture properties between the original and reconstructed halves, adds a quantitative dimension to the evaluation, shedding light on the nuances of the reconstruction process.

However, challenges in the preprocessing stage as preprocessing techniques were attempted, proved to be challenging or unsuccessful, those are as follows:

<u>1. Variable Duration and Aggregate Consecutive Stimuli:</u>
The attempt to handle variable stimulus durations and aggregate consecutive stimuli faced challenges. Additional data from the .mat file indicating stimulus durations was required, and grouping stimuli based on either total duration or a fixed number was not successfully implemented.

<u>2. Determine Group Size and Segment iEEG Data:</u>
Challenges were encountered in determining the optimal group size for consecutive stimuli and segmenting the iEEG data accordingly. Defining a suitable target segment duration and grouping stimuli based on it proved to be a complex task.

<u>3. Composite Visual Stimuli:</u>
Creating a composite representation for each group of visual stimuli, whether through concatenation or averaging, faced difficulties in implementation. The challenge lies in deciding the most effective method for combining individual stimuli images into a cohesive representation.

While these preprocessing steps faced challenges, addressing them could potentially enhance the reconstruction process. These challenges underscore the complexities of working with iEEG data and emphasise the importance of refining preprocessing techniques for more robust model training.

The project's interdisciplinary nature, spanning neuroscience, signal processing, and machine learning, contributes to the broader understanding of how neural patterns correlate with visual stimuli. While successful in certain aspects, it also points towards avenues for improvement, including refining preprocessing methodologies, exploring alternative model architectures, and potentially incorporating domain knowledge from neuroscience.

In conclusion, this project not only advances our understanding of decoding visual information from iEEG data but also sets the stage for future explorations in the intersection of brain-computer interfaces and neuroscience. The challenges encountered serve as valuable lessons for the scientific community, urging further collaboration and innovation in the pursuit of unravelling the complexities of the human brain.

REFERENCES

1. R. V. Rullen & L. Reddy (2018) - Reconstructing faces from fMRI patterns using deep generative neural networks
2. S. Nishimoto et al. (2011) - Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies
3. K. N. Kay et al. (2008) - Identifying natural images from human brain activity
4. M. Chelaru et al. (2016) - Reactivation of visual-evoked activity in human cortical networks
5. J. Dantzker (2006) - Bursting on the Scene: How Thalamic Neurons Grab Your Attention
6. T. Grootswagers et al. (2016) - Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data
7. F. Magnabosco & O. Hauk (2023) - Decoding semantics from dynamic brain activation patterns: From trials to task in EEG/MEG source space.
8. Dataset used: https://openneuro.org/datasets/ds004194/versions/1.0.1