Analysis of stimulus-driven EEG Activity Patterns in Humans

Biswadeep Chakraborty
Electrical and Computer Engg.
Georgia Institute of Technology

Lisa Meyer-Baese
Biomedical Engg.
Georgia Institute of Technology

Robert Nikolai
Biomedical Engg.
Georgia Institute of Technology

Jaxon Sommers
Biomedical Innovation and Development
Georgia Institute of Technology

Anni Zhou
Electrical and Computer Engg.
Georgia Institute of Technology

Abstract - During visual processing, the brain communicates through tangled network connections and compounded signal structures. However, we know that it is a trivial and quick task to recall the name of our favorite actors. To explain how the brain can pull off almost instantaneous face classification and to build off previous work in dimensionality reduction in the brain, we aim to model the brain's processing system in terms of optimized hardware and signals. Using methods including PCA we explore how the whole brain responds to complexity and how the frequency patterns of brain waves may adjust to complexity. We demonstrate that whole-brain responses to faces may increase in complexity with increased stimulus complexity. Next, we show that different types of images may be processed by distinct signal frequencies. Finally using a deep learning approach, we determined the regions of the outer cortex that are most important to face classification include the occipital and temporal cortex.

Index Terms - Face Classification, Dimensionality Reduction, Deep Learning, EEG, PCA

I. INTRODUCTION

To understand the fundamental principles of the human visual system, it is crucial to know how visual signals are propagated, organized, and processed. It is known that the brain processes information in a hierarchy of regions along its surface, ranging from "lower" areas that do basic parsing of incoming sensations to "higher" executive regions. Communication between these regions can be defined as signal propagation that produces a change in the representation by a recipient brain area. Effective inter-area communication relies on anatomical connections as well as physiological coordination on a millisecond time scale.

Across the cortex, the propagating signal is commonly studied in 4 frequency bands: theta (2-8 Hz), alpha (8-12 Hz), beta (12-35 Hz), and gamma (>35 Hz). Specific frequency bands of invasively and noninvasively recorded electrical activity have been shown to reflect aspects of visual stimulation such as categorical distinctions [1] or low-level stimulus features [2]. Given the ubiquity of these rhythms, it is important to know how they compare across cortical areas.

Across cortical areas, inter-area information transmission has been assessed using coherence measures across visual cortex V1- V2 [3] and by the likelihood of spikes in a recipient area given the state of a source area [4]. Studies have also investigated the timing (latencies) of signal propagation in the visual cortex and through the visual pathways [5], [6]. Such high-resolution studies have quantified the classical temporal hierarchy of information flow through the visual pathway. However, single-unit recordings may not be sufficient to encode the information flow on a larger scale.

Noninvasive methods of evoked potentials are valuable for detecting the temporal profiles of visual processing at a larger scale. Among the noninvasive methods, electroencephalogram (EEG) is the most practical method due to its fine temporal resolution. To aid in interpreting the outputs of collected EEG data, dimensionality reduction can be applied. When doing so it is important to vary the inputs to a brain area and ask whether the output of dimensionality reduction changes sensibly. Here we apply dimensionality reduction to both cortex EEG data and to channels corresponding to the primary visual cortex (V1) to ask two questions relating to changes in signal propagation across the cortex. First, how is the whole brain response related to stimulus complexity? Second, what frequency bands in V1 encode stimulus-related changes?

II. METHODS

A. Data

The dataset used here was obtained from a public EEG/MEG/fMRI dataset by Wakeman, D.G. & Henson, R.N. [7]. We are utilizing only the functional EEG data sampled at 1101Hz. This data was collected with a 74 channel headpiece. Two channels were identified to correspond to V1. The data were recorded from 16 human subjects performing a simple perceptual task involving classifying images of human faces. The images were of familiar (trigger = 5,6,7), unfamiliar (trigger = 13,14,15), or scrambled faces (trigger = 17,18,19). Familiar faces were selected from famous individuals whom the subject group can reasonably be expected to know. Each image (originally 162x128 pixels) was represented in a 20,736-dimensional space where each axis corresponds to the average intensity of one-pixel block. Continuous recording sessions consist of 150 stimuli responses with each class appearing 50 times. These sessions are divided into event epochs, with data in a 1-second window starting at stimulus onset yielding 1101 time points. In a session, epochs are grouped by channel and stimulus class and then averaged across its time series. The data is not averaged between recording sessions or subjects so we may compare variance and reproducibility. Raw EEG data were processed with the MNE python toolbox (mne.Epochs) [8] which preprocesses and extracts epochs from the raw instances.

B. PCA

All data were normalized and then evaluated with principal component analysis. Using the Scikit-learn python library, we computed the number of basis patterns needed to explain 90% of the variance in the recorded activity or visual stimuli.

C. V1 Frequency Analysis

An additional dataset is created to explore how V1 uses its signal frequency space. Prior to averaging the EEG signals by group, the time series were decomposed into separate frequency bands using hamming window filters. The chosen bands correspond to the aforementioned Gamma, Beta, Alpha, and

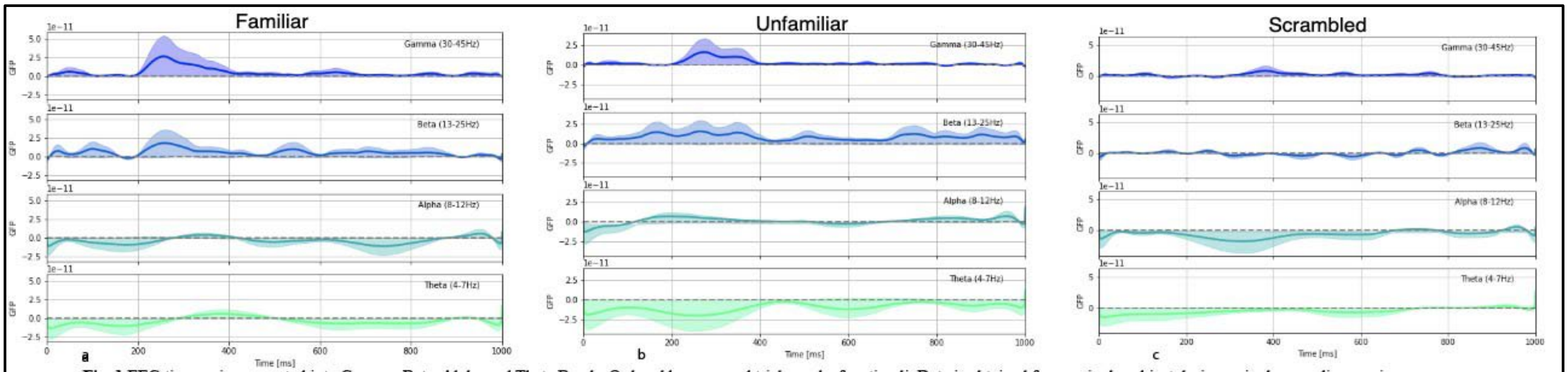

Fig. 1 EEG time series separated into Gamma, Beta, Alpha and Theta Bands. Ordered by averaged trial epochs for stimuli. Data is obtained from a single subject during a single recording session.

Theta bands. After filtering, the time series were averaged in a group of its frequency band, channel, and stimulus type (Fig. 1).

D. Filters

Testing accuracies were compared for three different filters. We used a high pass filter with a cutoff frequency of 1 Hz and a bandpass filter with a lower cut-off frequency of 1 Hz and an upper cutoff of 40 Hz. We also use a novel filter called the Neural Power Amplifier (NPA) filter [9] that combines filters with logarithmic and Gaussian magnitude responses.

E. Deep Learning Model

We used two different models to classify familiar, unfamiliar, and scrambled faces: a Long Short Term Memory (LSTM) network and a Multivariate Attention LSTM-FCN model. A graphical representation of the LSTM cell block is given in Fig. 2a. The complete network for the time series classification is given in Fig. 4b. Temporal convolutions have proven to be an effective learning model for time series classification problems [12]. Since the EEG data is a multivariate time series data, we use the multivariate attention-based LSTM-FCNN model for the classification. FCNs consisting of temporal convolutions are generally used to extract features. In the MALSTM-FCN model, the fully convolutional block is augmented by an LSTM block followed by dropout as shown in Fig. 2c.

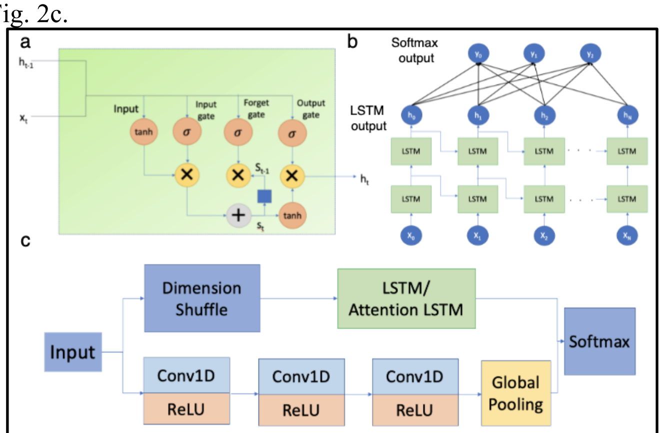

Fig 2 Model Architecture A: Cell diagram of each LSTM network consisting of the input, input gate, forget gate and the output gate. B: The LSTM model architecture. C: The Multivariate Attention LSTM-FCN Model using Attention LSTM

F. Feature Ranking and Shapley Values

The features extracted by the deep learning models are used for the classification of the three types of images. To get an analytical evaluation of these features, we used the SHAP method [13] to estimate the contributions of different features in the classification. SHAP uses the concept of Shapley values from game theory to explain the predictions of a machine learning model in a similar way as computing the contributions of players in a game.

III. RESULTS

A. Whole Brain

Taking the data from all 74 EEG channels, we wanted to see whether there are qualitative differences between the principal components of the different stimuli. As shown in Fig. 2b, the top principal component (PC) for each stimulus that explains between 35-49% of the variance in the data contains mainly low-frequency information. Higher frequency components can be seen in the 3rd PC which explains less variance.

To quantify the similarity of the identified patterns we used an aggregation method as done by Cowley et al [10]. First, as a baseline, we tested to see the visual stimuli themselves to see the extent to which they reside in the same dimensions in pixel space. We compared the dimensions of individual stimuli (Fig. 3c blue dots) and combinations of them (Fig. 3c orange and purple dots).

If stimuli reside in overlapping dimensions, then the resulting dimensionality would be the maximum of the dimensionalities for the individual stimuli. Alternatively, if the stimuli reside in orthogonal dimensions, then the resulting dimensionality would be the sum of individual stimuli.

The main observation from Fig. 3c is the dimensions occupied by the familiar faces have some overlap with those for unfamiliar and scrambled. Interestingly, the scrambled images were not much more complex in the pixel space but required more dimensions in the EEG data. We did the same analysis using the subjects' EEG data (results followed a similar trend for other subjects). Here we found that familiar and unfamiliar faces share more common PCs, while the PCs for familiar faces are fully a subset of unfamiliar and scrambled faces.

B. V1 Single Channel

This analysis was done with the data split into frequency bands. This time, only V1 is considered so the time series data for the two channels relating to this region were averaged. At this point, averaged signals for each stimulus group are taken to be four-dimensional time-series with frequency bands as dimensions. To explore how the

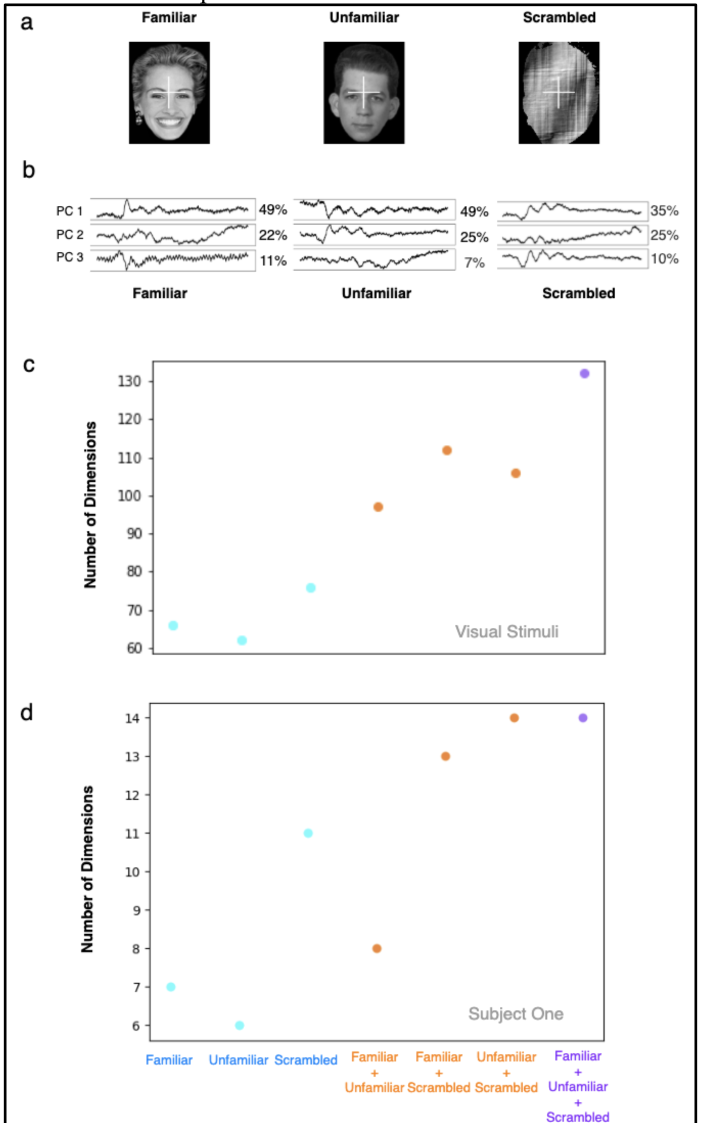

Fig 3 Similarity of principal components across stimuli A: Example images for each category that were presented to participants. B: Principal components describing the largest percent variance for the whole brain response to familiar, unfamiliar and scrambled images in subject 1. Percentages denote the percent variance explained by each pattern. C: Dimensionality of visual stimuli for individual images (blue), combinations of the two (orange) and all three (purple). D: Dimensionality of EEG response for individual and combinations of stimuli

dimensionality of the frequency space changes across the stimuli, we determine the number of dimensions needed to explain 90% of the variance using PCA (Fig. 4). The Gamma, Beta, Alpha, and Theta bands were chosen for this analysis because brain waves are widely thought to be organized into these groups of frequency or frequency patterns for similar functions. These frequency patterns are specifically considered as related to conscious thought. From the PCA analysis, we cannot make an interpretation of which bands are important for these stimulus tasks. Rather, our goal is to determine if the frequency patterns are the same or distinct for varying stimuli. The blue dots of Fig. 4 consistently demonstrate that the complexity of the frequency space is consistently low (<4) between the responses in the different stimulus groups. To compare the frequency patterns between the stimulus groups, all three of the stimulus group's time series data was aggregated in combinations of two or three groups and then run with PCA. The same method from the whole brain analysis applies here. As seen in Fig. 4, for two combined stimuli, the dimensions hover between 3 and 4 principal components indicating little to no overlap. For the combination of 3 stimuli, the dimension hovers between 4 and 5 dimensions, which more strongly suggests at least some overlap. This might indicate that the frequency patterns of two stimuli represent enough of the total frequency space used for face processing to encompass the third.

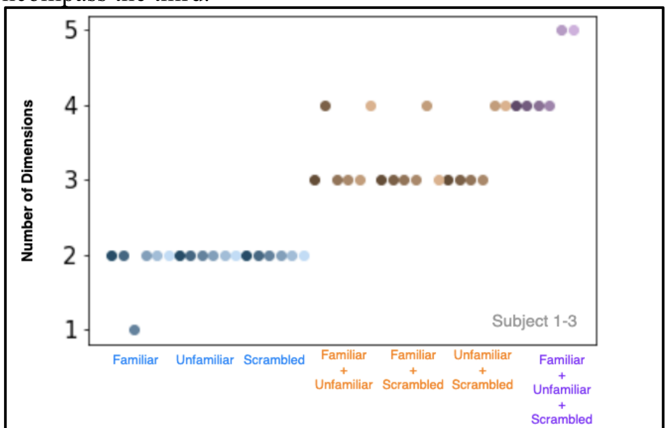

Fig 4 Dimensionality of frequency spectrum for individual image classes (blue) and combinations of two (yellow brown) or three (purple) classes. For each class or combination, there are 6 sets of data (dark to light dots). This is obtained from two recording sessions from three subjects each.

C. *Deep Learning Model*

We trained the models described in Section II on the OpenNeuro dataset to distinguish between input stimuli. We used the SHAP methodology to rank the important features (the EEG channels) and compared these to their anatomical regions. After evaluating the models we get the following observations:

(a) The basic LSTM model, though a basic version, then ran for 4 folds over the 16 subjects in the datasets, gives an overall accuracy of 34.7%. On the other hand, the MALSTM-FCN model which improves on the basic LSTM module achieves an overall performance of 69.4% over the same dataset.

(b) We rank the features from the MALSTM_FCN model using the SHAP methodology to see the classification ability of each of the channels used. We see that the top electrodes according to the feature learning algorithm are located primarily along the occipital lobe and the temporal lobe of the brain. This observation is aligned to the prior works [7], [14]. This is in line with previous studies, which have linked this presentation frequency to early visual cortex responses [15], [16].

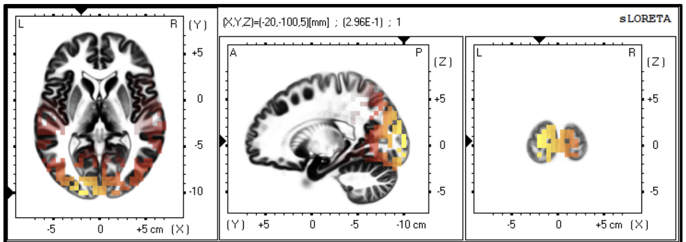

Fig 5 sLORETA solutions showing the regions of the brain corresponding to the top 20 electrodes with the highest SHAP values. These regions of the brain help in the better classification of the familiar, unfamiliar and scrambled faces for the MALSTM-FCN model.

IV. DISCUSSION

Our first goal was to investigate how the whole brain response relates to stimulus complexity. To do this, we applied PCA to trial-averaged functional EEG data in response to different classes of visual stimuli. We found that the PCs for familiar, unfamiliar, and scrambled faces shared some overlap. Overall this suggests that cortex wide processing of visual information is capable of expressing a finite repertoire of principal components and that a subset of those components is employed for any given stimulus. We note that even though the EEG responses to different stimuli occupy some similar dimensions, this does not imply that the responses activate the same regions of the PC subspace. The activity we measure may covary along the same dimensions but be centered at different locations in the functional EEG space. Future work can look into whether this activity is centered in the brain.

Our second goal was to understand the frequency patterns that V1 uses to process images of varying complexities. We performed this analysis by separating frequencies into smaller bands rather than individual components or features to create a simple and interpretable model. From our band-level analysis, we found little overlap between the frequency pattern components of different stimuli responses. However, in comparison to the trends seen in Fig. 3c, the dimensions for image combinations increased similarly to the combinations of the responses they invoke. Perhaps the number of frequency components used to explain an image scale up with image complexity. In the future, this could be further studied by decomposing the signal into more basic frequency components using a wavelet transform method.

From our model and ranked features we noted that the top 20 channels needed for feature classification come from V1. This suggests the stimulus-driven patterns are best discerned in V1 and not in other areas along the visual pathway. For future work, it would be interesting to connect the results from the whole-brain PCA to see what specific regions are similar/different for the varying stimuli.

V. CONCLUSION

To aid in the understanding of stimulus-driven EEG activity patterns, we chose to break down whole-brain EEG using dimensionality reduction. We found that for whole-brain activity, responses to different stimuli occupy similar dimensions of the EEG space. In our signal analysis for V1, we found that different types of stimuli may invoke different frequency patterns at lower dimensions. For comparison, we applied a similar approach using a deep convolutional network to assess what anatomical regions were most relevant for stimuli classification.

Our work is a preliminary analysis to assess the dimensionality and similarity of PCs for stimulus-driven EEG activity.

VI. References

[1] Gonzalez Andino, S.L., Grave de Peralta, R., Khateb, A., Pegna, A.J., Thut, G., Landis, T., 2007. A glimpse into your vision. *Hum Brain Mapp* 28, 614–624. doi:10.1002/hbm.20302

[2] Besserve, M., Lowe, S.C., Logothetis, N.K., Schölkopf, B., Panzeri, S., 2015. Shifts of Gamma Phase across Primary Visual Cortical Sites Reflect Dynamic Stimulus-Modulated Information Transfer. *PLoS Biol* 13, e1002257. doi:10.1371/journal.pbio.1002257

[3] X. Jia, S. Tanabe, and A. Kohn, "Gamma and the Coordination of Spiking Activity in Early Visual Cortex," *Neuron*, vol. 77, no. 4, pp. 762–774, 2013.

[4] A. Zandvakili and A. Kohn, "Coordinated Neuronal Activity Enhances Corticocortical Communication," *Neuron*, vol. 87, no. 4, pp. 827–839, 2015.

[5] Nowak L, Bullier J. The timing of information transfer in the visual system. In: Rockland KS, Kaas JH, Peters A, editors. Extrastriate cortex in primates. New York: Springer; 1997. pp. 205–241

[6] Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG. Signal timing across the macaque visual system. J Neurophysiol. 1998;79:3272–3278.

[7] Wakeman, D.G. & Henson, R.N. (2015). A multi-subject, multi-modal human neuroimaging dataset. Sci. Data 2:150001 doi: 10.1038/sdata.2015.1

[8] A. Gramfort, M. Luessi, E. Larson, D. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, M. Hämäläinen, MEG and EEG data analysis with MNE-Python, *Frontiers in Neuroscience*, Volume 7, 2013, ISSN 1662-453X

[9] Doyle, J.A., Toussaint, P.J. and Evans, A.C., 2019. Amplifying the Neural Power Spectrum. *BioRxiv*, p.659268.

[10] B. R. Cowley, M. A. Smith, A. Kohn, and B. M. Yu, "Stimulus-Driven Population Activity Patterns in Macaque Primary Visual Cortex," *PLOS Computational Biology*, vol. 12, no. 12, 2016.

[11] Karim, F., Majumdar, S., Darabi, H. and Chen, S., 2017. LSTM fully convolutional networks for time series classification. IEEE access, 6, pp.1662-1669.

[12] Wang, Z., Yan, W. and Oates, T., 2017, May. Time series classification from scratch with deep neural networks: A strong baseline. In 2017 International joint conference on neural networks (IJCNN) (pp. 1578-1585). IEEE.

[13] Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S.I., 2020. From local explanations to global understanding with explainable AI for trees. Nature machine intelligence, 2(1), pp.56-67.

[14] Collins, E., Robinson, A.K. and Behrmann, M., 2018. Distinct neural processes for the perception of familiar versus unfamiliar faces along the visual hierarchy revealed by EEG. NeuroImage, 181, pp.120-131.

[15] Müller, M.M., Andersen, S., Trujillo, N.J., Valdes-Sosa, P., Malinowski, P. and Hillyard, S.A., 2006. Feature-selective attention enhances color signals in early visual areas of the human brain. Proceedings of the National Academy of Sciences, 103(38), pp.14250-14254.

[16] Robinson, A.K., Venkatesh, P., Boring, M.J., Tarr, M.J., Grover, P. and Behrmann, M., 2017. Very high density EEG elucidates spatiotemporal aspects of early visual processing. Scientific reports, 7(1), pp.1-11.

PROJECT CONTRIBUTIONS

Biswadeep Chakraborty: Data processing, coding and training deep learning model, writing report. Lisa Meyer-Baese: Data loading and preprocessing, image registration, whole brain PCA analysis, writing report. Robert Nikolai: Stimulus image analysis and PCAs, frequency and signals processing and analysis, frequency band PCA analysis, writing the report. Jaxon Sommers: Organized team, project scope, report formatting, research, writing report. Anni Zhou: Data loading, deep learning model debugging, writing report.