Brainwaves and Emotions: Mapping the Mind with Machine Learning

Joshua Green, Bach Nguyen, and Bond Leung
Department of Computer Science
University of Missouri
Columbia, MO 65211
{bllndd, jwgktd, bxnwx2}@umsystem.edu

1 About our Datasets

1.1 EEG Data

The overarching goal of SEED-IV was to provide a rich, multimodal dataset that supports the development and evaluation of emotion recognition algorithms. It aims to **advance understanding of how emotions manifest in the brain and ocular behavior**, and to enable robust, real-time affective computing systems. Particularly using EEG and eye movement data. The core premise was to collect synchronized and physiological data while participants were exposed to emotional stimuli. Exploring how brain activity and eye behavior reflect emotional states.

1.2 MRI Data

The study "**Localizing Brain Function Based on Full Multivariate Activity Patterns (MVAP): The Case of Visual Perception and Emotion Decoding**" successfully showcased that brain-wide MVPA can be a viable tool not just for classification but for functional brain mapping. Valuable in detecting subtle and distributed patterns associated with emotional perception. The fMRI data were collected from 16 subjects performing a face perception task involving happy, sad, angry, neutral, and scrambled faces.

Noting that happiness was the only emotion that consistently yielded statistically significant, localized brain patterns. While sadness and anger did not produce clear or consistent activation patterns across subjects.

2 EEG Experiments

2.1 Supervised

We employed two supervised learning models—Support Vector Machine (SVM) with a linear kernel and a feed-forward Multi-Layer Perceptron (MLP)—to classify emotional states using EEG data, evaluating them through accuracy metrics and confusion matrices. To enhance feature engineering, we applied Principal Component Analysis (PCA) for dimensionality reduction, testing models in three configurations: without PCA, with standard PCA, and with PCA optimized via validation. All models were trained on the same dataset using grid search and cross-validation for fair comparison

2.2 Unsupervised

In addition to the supervised classification model, an unsupervised learning component was incorporated to explore whether natural groupings in the EEG data aligned with known emotional categories. This was achieved using a combination of Principal Component Analysis (PCA) for dimensionality reduction, t-distributed Stochastic Neighbor Embedding (t-SNE) for visualization, and K-Means clustering to detect latent structure. The results present coherent clusters indicating that EEG signals encode sufficient discriminative information for emotion differentiation, likely stemming from the neural dynamics evoked by the emotion-eliciting film stimuli.

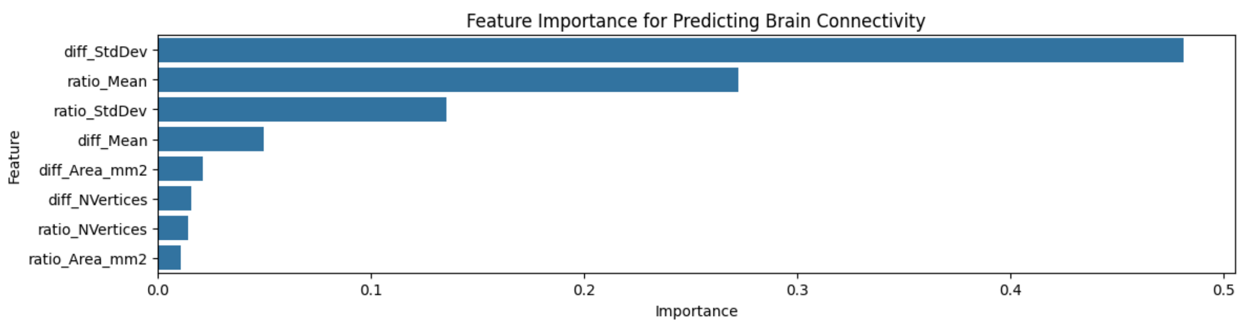
3 MRI Data Processing

3.1 Freesurfer

To extract functional Magnetic Resonance Imaging (fMRI) surface projections and anatomical statistics for analysis, several shell scripts were employed to process subject-level data using FreeSurfer and FSL utilities (a comprehensive suite, software library of analysis tools for functional and structural brain imaging data). To obtain

surface-mapped BOLD data and quantitative brain metrics for use in machine learning models.
Script 1 (rec_vol2surf.sh) automated the recon-all and mri_vol2surf processes for all subjects located in an input directory. Not all subjects were completed successfully due to runtime. Script 2 (vol2surf_run.sh) was created for the subjects that failed due to timeouts, which applies mri_vol2surf to subjects that have been run through recon-all individually. Script 3 (preprocess_bold.sh) handled BOLD fMRI preprocessed images using FSL tools. It included slice timing correction “slicetimer”, which ended up being used for unsupervised learning on the fMRI data. The volumes were then projected to the fsaverage surface for both hemispheres via mri_vol2surf and subsequently extracted statistics using mri_segstats. Script 4 (extract_stats2.sh) leveraged FreeSurfer’s aparcstats2table and asegstats2table to gather measures like cortical thickness, volume, surface area, and curvature.

3.2 Feature Analysis
We analyzed structural brain measurements and connectivity matrices across subjects to study brain network organization. Pairwise regional features were used to predict connectivity strength, and PCA was applied to visualize subject clustering. Graph-theoretical metrics—such as density, clustering coefficients, and centrality—quantified network topology. Random forest regression identified key predictive features, while regional connectivity scores highlighted highly connected areas. Visualizations included PCA plots, brain network graphs colored by centrality, feature importance charts, and regional connectivity bar plots.



4 Results and discussion

4.1 EEG Results
SVM demonstrated stronger recall for Neutral and Happy emotions, while MLP excelled in detecting Fear, suggesting an advantage in capturing nuanced emotional features. Applying PCA preserved SVM performance and significantly improved MLP accuracy to 69%, enhancing precision and recall for Neutral and Fear. However, combining PCA with hyperparameter tuning via GridSearch led to declines likely due to overfitting. Among all configurations, MLP with PCA (MLPa) yielded the best and most balanced results, closely aligning with the 70.33% accuracy reported in prior SVM-based EEG studies using differential entropy features. Unlike Power Spectral Density, differential entropy captures signal complexity more effectively. Notably, reducing variance below 99.9% during PCA led to accuracy drops, emphasizing that even minimal variance components hold valuable information. These findings underscore the critical role of preprocessing and tailored model tuning in enhancing emotion recognition from EEG data.

4.2 MRI Results
The strongest emotion decoding performance came from **happy**, with high accuracy and statistically significant anatomical localization, matching the MRI data study. Both finding happiness is most decodable. **Angry** was second with only being misclassified once as **blank**, which is still inline with the study, which found anger to be a bit harder to decode. **Sad** and **neutral** had some confusion with each other, with inconsistent anatomical patterns and lower classification performance. **Scrambled** and **blank** had small sample sizes but show mutual confusion, possibly due to visual similarity or reduced task engagement.

4.3 Next Steps
After observing the unsupervised learning for both EEG and fMRI data, a meaningful similarity emerged in both modalities, reflecting emotion-based separability and clustering consistency. Strongly valenced emotions like *happy*

and *sad* showed mostly clear separations, while *neutral* remained less distinct. This cross-modal consistency validates the robustness of emotion-specific brain signatures and supports the multimodal approach to affective state decoding in our project.

To further build on these findings, we propose testing a sequential model. A Long Short-Term Memory (LSTM) network on the chess dataset. Unlike static classifiers, an LSTM can capture temporal dependencies and evolving patterns within the data, making it suitable for identifying **latent or transitional emotional states** during cognitive emotion labels and exploring how emotions fluctuate over time in response to task complexity, stress, or decision-making under pressure. Aligning with our broader goal of dynamic emotion modeling across neural modalities.

Acknowledgment

We gratefully acknowledge the use of fMRI emotion data provided in the study by Gu et al. (2021), *"Shared affective experience across participants revealed by dynamic facial expressions and self-reported emotions in naturalistic fMRI."* This dataset was accessed via bioRxiv and played a critical role in supporting the multimodal emotion recognition analysis in our research. We thank the authors for making their work publicly available and contributing to open neuroscience.

References

[1] Wei-Long Zheng, Wei Liu, Yifei Lu, Bao-Liang Lu, and Andrzej Cichocki, EmotionMeter: A Multimodal Framework for Recognizing Human Emotions. IEEE Transactions on Cybernetics, Volume: 49, Issue: 3, March 2019, Pages: 1110-1122, DOI: 10.1109/TCYB.2018.2797176.