

# Toward Scalable Multimodal Neuroimaging Analysis

## A FABRIC-Automated Pipeline for Emotion Recognition and Brain Efficiency Mapping

### Source

- [Brain Efficiency Comparison Between Chess Masters and Amateurs Using Structural MRI Metrics](#)
- [EEG and fMRI Emotion Recognition Using Machine Learning and Automated Pipelines](#)

By: **Joshua Green, Bach Nguyen**

Department of Computer Science, University of Missouri, Columbia, MO 65211

### Abstract

Neuroscience research increasingly relies on scalable analysis pipelines to handle high-dimensional multimodal data. This work presents an automated system for processing EEG and MRI signals to support emotion recognition and brain efficiency analysis. We integrate supervised and unsupervised machine learning models with scripted pipelines executed on the FABRIC testbed. EEG and fMRI data are used for classifying emotions, while structural MRI metrics are employed to compare cortical organization in expert versus non-expert cognitive populations (e.g., chess masters vs. controls). Results show effective emotion clustering and significant anatomical differences across groups, demonstrating the potential for automated pipelines in large-scale cognitive neuroscience.

---

### 1. Introduction

Affective state recognition and cognitive performance benchmarking are crucial in neuroscience, education, and adaptive systems. However, these tasks require processing complex datasets, including electroencephalography (EEG) and magnetic resonance imaging (MRI). This project contributes a scalable framework that (1) classifies emotional states using multimodal signals, (2) maps structural brain differences tied to cognitive expertise, and (3) automates preprocessing through Bash scripting and distributed testbed execution.

Despite efforts to integrate EEG and fMRI, one core challenge we encountered was the incompatibility of spatial and temporal dimensionality. EEG recordings are fundamentally 3D (channel  $\times$  time  $\times$  trial), while fMRI signals often reside in 4D (voxel  $\times$  voxel  $\times$  slice  $\times$  time). As a

result, we could not directly align time-series predictions or apply shared clustering techniques across both modalities. For future work, techniques like canonical correlation analysis (CCA) or shared embedding spaces may help resolve these incompatibilities.

---

## 2. Data Sources and Preprocessing

**EEG Emotion Dataset:** SEED-IV EEG data includes recordings from 15 participants across 72 emotional film stimuli. After re-referencing, filtering (1–75 Hz), and segmentation (4s epochs), Differential Entropy (DE), Power Spectral Density (PSD), and statistical features were extracted.

**fMRI Emotion Dataset:** fMRI volumes from face-perception tasks (happy, sad, neutral, angry) were processed using FreeSurfer. Volumes were mapped to cortical surfaces using `mri_vol2surf` and analyzed using `mri_segstats`.

**sMRI Brain Efficiency Dataset:** Structural MRI scans from 58 subjects (29 chess masters, 29 controls) were processed with `recon-all`, `aparcstats2table`, and `asegstats2table` to extract cortical thickness and subcortical volume features.

**Automation Tools:** Custom Bash scripts were hosted on GitHub and executed across nodes on the FABRIC testbed. These included:

- EEG: Filtering, re-referencing, feature extraction
  - MRI: `rec_vol2surf.sh`, `preprocess_bold.sh`, `extract_stats2.sh`, `vol2surf_run.sh`
- 

## 3. Machine Learning Models

**EEG/fMRI Emotion Recognition:**

- Supervised: Support Vector Machine (SVM), Multi-Layer Perceptron (MLP)
- Unsupervised: PCA, t-SNE, K-Means clustering

**sMRI Brain Efficiency Comparison:**

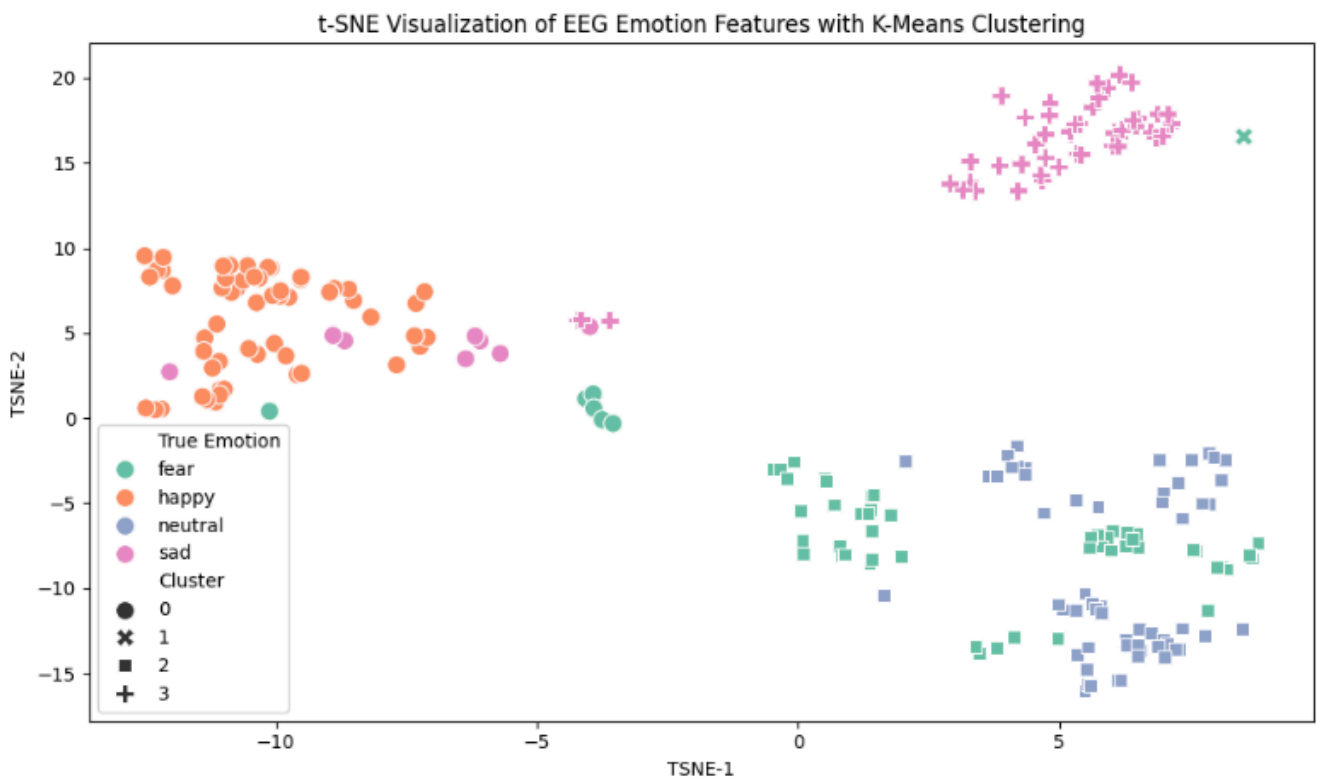
- Welch's t-test on key features
  - PCA and KMeans for visual and statistical separation
- 

## 4. Results

**Emotion Classification:**

- MLP with PCA yielded 69% EEG accuracy, especially for happy and neutral states.
- fMRI showed strong anatomical activation patterns for happy and fear stimuli.
- t-SNE projections revealed distinct emotion clusters from EEG-derived features.

**Figure 1.** t-SNE projection of EEG emotion features showing distinct clustering of happy, sad, fear, and neutral states.



**Brain Efficiency Differences:**

- Cortical thickness in superior parietal and precuneus regions significantly higher in chess masters ( $p < 0.001$ ).
- BrainSegVol-to-eTIV ratio also significantly elevated ( $p = 0.0005$ ), suggesting higher structural efficiency.
- PCA on sMRI features showed partial group separation, supported by KMeans clusters.

**Table 1.** Top Structural MRI Features Differentiating Chess Masters and Controls

Feature	Mean Diff	p-value
rh_superiorparietal_thickness	+0.0959	0.00005
lh_supramarginal_thickness	+0.1106	0.00020
lh_precuneus_thickness	+0.0984	0.00054

Feature	Mean Diff	p-value
BrainSegVol-to-eTIV	+0.0036	0.00054
lh_paracentral_thickness	+0.1542	0.00066

---

## 5. Discussion

This integrated system bridges emotional decoding and cognitive performance benchmarking using a shared neuroimaging pipeline. EEG’s temporal sensitivity enabled dynamic emotion tracking, while fMRI and sMRI provided spatial context and anatomical markers. Automation via FABRIC scripts ensured reproducibility, parallelization, and scalability across neuroimaging workflows. The combination of dimensionality reduction, feature engineering, and ML models yielded biologically meaningful insights.

The dimensionality challenge between EEG and fMRI remains a barrier to full cross-modal fusion. Future approaches could explore encoding each modality into a shared latent space or aligning them via temporal synchronization using surrogate signals.

---

## 6. Conclusion

Our pipeline advances neuroimaging workflows by unifying emotion recognition and brain efficiency modeling in a reproducible and scalable architecture. By leveraging FABRIC, FreeSurfer, and modern ML techniques, this work supports high-throughput brain-behavior mapping—crucial for affective computing, skill assessment, and future brain-inspired networks.

---

## 7. Future Work

The current work can be extended by introducing dynamic sequence modeling for EEG data using Long Short-Term Memory (LSTM) networks. This would allow real-time temporal pattern detection across emotional states. For structural and functional MRI, applying graph neural networks (GNNs) on brain connectivity graphs could reveal complex regional interactions associated with cognition and emotion.

To overcome the dimensionality barrier between EEG and fMRI, the aim is to explore multi-view learning techniques such as shared autoencoders, canonical correlation analysis (CCA), or

contrastive representation learning. These approaches could create shared embeddings that unify temporal EEG dynamics and spatial fMRI patterns.

Furthermore, containerizing the entire preprocessing pipeline (e.g., Docker or Singularity) and deploying it through a distributed scheduler would ensure even greater reproducibility and scalability. Real-time systems and cross-population generalization, including clinical or high-performance user groups, remain important areas for exploration.

---

**References** [1] Wei-Long Zheng et al., EmotionMeter: A Multimodal Framework for Recognizing Human Emotions, *IEEE Trans. Cybernetics*, 2019.

[2] Gu et al., Shared affective experience in naturalistic fMRI, *bioRxiv*, 2021.

[3] SEED-IV Dataset: <https://bcmi.sjtu.edu.cn/~seed/>

[4] FreeSurfer: <https://surfer.nmr.mgh.harvard.edu/>

[5] GitHub Pipeline Repository:

[https://github.com/JoshGreen000/4001\\_CI\\_Bio\\_Eng\\_Final\\_Project](https://github.com/JoshGreen000/4001_CI_Bio_Eng_Final_Project)

[6] SCinet INDIS 2025 CFP: <https://scinet.supercomputing.org/community/indis/indis2025/>