Neural Decoding of Visual Stimuli

Using Mice Local Field Potential Power Spectral Density

Gabriele Paganelli & Phillip Angelos • University of Padova / Boston University & Boston University

78–79%	+10–15 pp	∼96k	8
Test Accuracy	vs. Linear Baseline	Training Trials	Stimulus Classes

Overview

We build a neural decoder that identifies which of 8 visual stimuli a mouse is viewing, using only brain signals recorded from its visual cortex. The model takes 1-second windows of Local Field Potential (LFP) data, extracts frequency-band power features, and classifies the stimulus via a GRU–Transformer hybrid. The approach is motivated by **translational neuroscience**: architectures validated on mouse LFP can inform future work decoding human EEG for clinical applications such as tracking cognitive states and neurological biomarkers.

Data

Source: Allen Institute Visual Coding Neuropixels dataset (7 recording sessions, multiple mice). Raw LFP is preprocessed by Allen: spatial/temporal downsampling, 0.1 Hz high-pass filtering, and re-referencing to remove common-mode noise.

Features: Welch's PSD is computed per channel and segmented into 4 frequency bands (*theta, alpha, beta, gamma*) across 10 time bins, yielding a (K=10, D=532) matrix per trial.

Class distribution: Highly imbalanced. Natural movies and Gabors have ∼21–24k trials each; dot motion and flashes have only 1–1.5k trials.

Natural movie (shuffled)	Natural movie (more repeats)
Drifting gratings (75 rep.)	Drifting gratings (contrast)
Gabor patches	Dot motion
Flashes	Spontaneous activity

Split: Stratified 70/30 train/test; features standardized using training-set statistics only.

Model Architecture

A **GRU–Transformer hybrid** exploits temporal structure in LFP bandpower:

1. **Input projection:** linear layer + LayerNorm, maps $D \to d_{\text{model}}$
2. **Transformer encoder:** self-attention over the K time bins, capturing long-range dependencies
3. **GRU layer:** refines representations with gated recurrence, adding inductive bias for sequential dynamics
4. **Classifier head:** LayerNorm → Linear → ReLU → softmax over 8 classes

GRU is preferred over LSTM for efficiency and reduced overfitting on minority classes. Training uses **Adam**, **class-weighted cross-entropy** (to handle imbalance), and augmentation (Gaussian noise + circular time shifts).

Results

Stimulus	Recall
Nat. movie shuffled	0.88–0.94
Gabors	0.80–0.91
Nat. movie repeats	0.68–0.73
Drifting grat. (75r)	0.69–0.76
Spontaneous	0.70–0.79
Drifting grat. (cont.)	0.54–0.65
Flashes	0.56–0.70
Dot motion	0.50–0.72

Per-class recall

The GRU-Transformer's **largest gains** appear on stimuli with distinctive temporal dynamics (drifting gratings, dot motion), confirming that the temporal evolution of bandpower, not static power alone, carries critical discriminative information.

Key Findings

- **Temporal modeling matters:** +10–15 pp over static baseline, consistently across all runs and sessions.
- **Confusion is structured:** errors occur between stimuli that are visually similar (e.g., the two drifting grating conditions), reflecting genuine neural similarity rather than random noise.
- **Imbalance is manageable:** class-weighted loss + augmentation substantially improves recall on minority classes (flashes, dot motion).
- **Translational relevance:** the full pipeline (PSD features, time binning, GRU-Transformer, class weighting) is directly compatible with human EEG analysis.

Technologies

Languages & libraries: Python, PyTorch, AllenSDK, NumPy, SciPy (Welch's PSD), scikit-learn.
Data: Allen Institute Neuropixels Visual Coding dataset.
Code: `github.com/GabriPaganelli/Mouse-LFP-Transformer`

Future work: multi-session transfer learning and cross-animal generalization; attention-based channel selection for cortical region importance; extension to raw LFP waveforms via end-to-end convolutional encoders.