



## Where are thy, O simple spike?

*A data-driven approach to brain region identification in electrophysiological recordings.*

### ABSTRACT

Accurate anatomical localization of Neuropixels recordings is essential for systems neuroscience but is limited by post-hoc histology and atlas registration, which are prone to observer bias and inter-subject variability. Here, we introduce a data-driven framework for classifying brain regions directly from electrophysiological signatures, reducing reliance on anatomical mapping. Using the Allen Brain Observatory Neuropixels Visual Coding dataset, spike-train data were processed into two complementary streams: (1) eight handcrafted temporal features (e.g., peak latency, firing rate, ISI covariance) and (2) full peri-stimulus time histograms (PSTHs). Classical classifiers were compared with Bayesian-optimized, two-stream deep learning architectures, including DMLP, LDA-Dense, CNN, TCN, and Transformer models. Classical approaches failed to generalize (weighted accuracy <0.19), while deep models achieved substantial gains. The Transformer performed best (weighted accuracy = 0.35; AUC = 0.86), particularly in thalamus (0.56). These results indicate that electrophysiological activity encodes partial anatomical structure. This framework may reduce reliance on post-mortem histology and limit invasive procedures in animal research.

**Casper van Laar—2440678**  
[casperdvanlaar@hotmail.com](mailto:casperdvanlaar@hotmail.com)  
University of Wolverhampton  
Msc artificial intelligence  
7CS078/UZ1: Data Science  
31/10/2025  
Word count: 3083

## PROBLEM STATEMENT:

Large-scale electrophysiological recordings provide a powerful means to study neural circuits, but accurately localizing recorded units remains challenging due to anatomical and histological variability. Neuropixels probes are high-density silicon electrodes capable of simultaneously recording hundreds to thousands of neurons across multiple brain regions at single-cell resolution. They offer an unprecedented view of neural population dynamics in awake, behaving animals (Steinmetz, 2021). These advances have transformed neuroscience by enabling large-scale, simultaneous spiking measurements, driving adoption of standardized frameworks such as the Allen Common Coordinate Framework (CCFv3) for spatial registration. However, current localization methods relying on post hoc histological reconstruction and registration to the averaged CCFv3 (Wang et al., 2020) introduce uncertainties that compromise accuracy and comparability across datasets.

The main sources of error are twofold: manually marking registration points and probe tracks introduces observer variability, and the CCFv3 itself is an averaged atlas that fails to capture natural variation in brain size, shape, and cytoarchitecture. As probes penetrate deeper, assigning fixed anatomical labels to channels in uniquely shaped brains becomes increasingly uncertain, often requiring subjective adjustments based on physiological features. This cumulative methodological and biological uncertainty limits confidence in precise electrode localization, particularly for deep structures, despite the sophistication and proven reliability of histological techniques (Kiernan, 2015). Considering that animal models are living beings, developing a data-driven registration framework directly from Neuropixels recordings represents not only a methodological but also a moral advance, enabling accurate localization without post-mortem histology.

This study explores a data-driven alternative: identifying and classifying brain areas corresponding to electrode channels without graphical or manual mapping. The approach rests on a neurophysiological principle: each brain region exhibits a unique electrophysiological signature (e.g., spike patterns, unit activity) in response to a standard stimulus, and responses from the same region across mice are more similar than responses from distinct regions within one animal (Tolossa et al., 2025; Kordecka et al., 2025; International Brain Laboratory, 2024). Based on this principle, brain regions can be classified directly from neural firing signatures, bypassing anatomical coordinates. Leveraging large-scale Neuropixels data, this work develops a machine learning framework to infer anatomical labels directly from electrophysiological activity.

# Dataset Source of dataset and citation pre-processing

## Source of Dataset

The dataset is publicly available through the AllenSDK and distributed in Neurodata Without Borders (NWB) format (Allen Institute for Brain Science, 2019).

The dataset originates from the Allen Brain Observatory Neuropixels Visual Coding Project (Allen Institute for Brain Science, 2019) and contains large-scale electrophysiological recordings using Neuropixels Phase 3A probes, measuring spiking and local field potential activity across cortical and subcortical regions in awake mice exposed to visual stimuli.

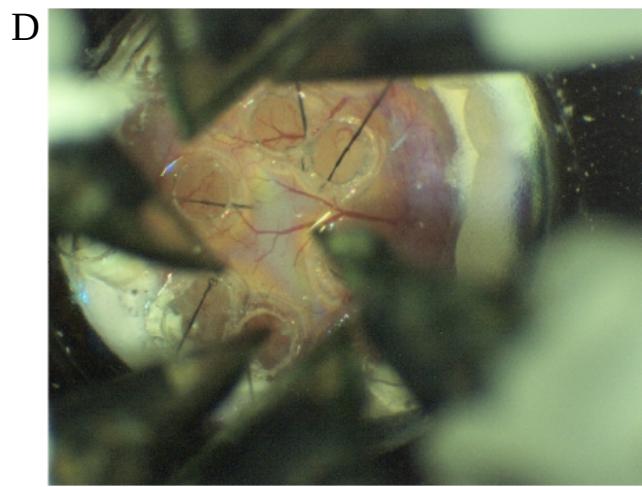
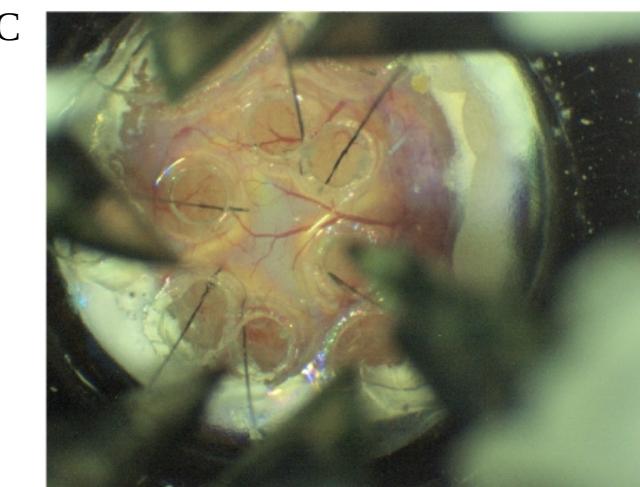
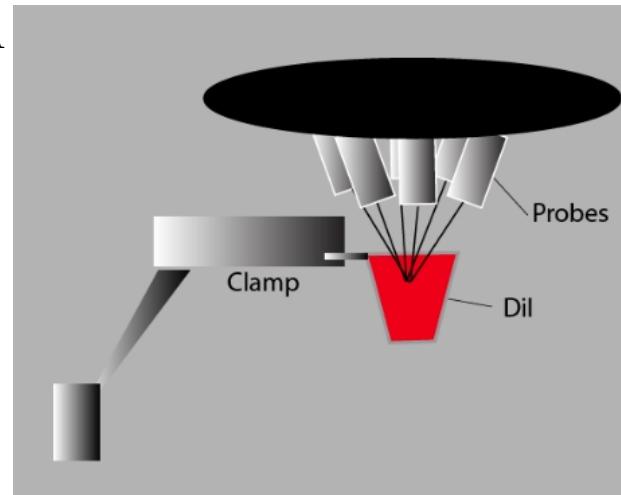
Each experiment used up to six probes, each with 384 active sites spanning primary visual cortex (V1Sp), higher visual areas (VISl, VISam, VISpm, VISrl), thalamic nuclei (LGd, LP), superior colliculus (SC), and hippocampal regions (CA1, CA3, DG). Visual stimuli included full-field flashes, drifting gratings, natural scenes, and movie clips, presented in standardized sequences for reproducibility.

Recordings were obtained from wild-type and transgenic Cre-driver mice (Tsien, 2016) and aligned to the Allen Mouse CCFv3 via postmortem histology, providing anatomical coordinates per channel. The dataset is publicly available through the AllenSDK and distributed in NWB format (Allen Institute for Brain Science, 2019).

## Data Acquisition

Each mouse was head-fixed and exposed to a visual display while Neuropixels probes recorded neural signals at high temporal resolution (fig 1).

- Sampling rates: 30 kHz (spike band) and 2.5 kHz (LFP band)
- Recording duration: 2–3 hours per session
- Stimuli: Flashes, gratings, natural images, and movies
- Regions: Cortex, thalamus and hippocampus. Divided into subareas.
- Behavioral signals: Pupil diameter, running speed
- Synchronization: Hardware-level photodiode timing for precise alignment of spikes and stimuli



*Figure 1: Recording setup.* (A) “Prior to recording, 6 probe shanks are submerged in a bath of CM-Dil to a depth of at least 4 mm. A plastic well was glued to a standard headframe and clamped into the headframe holder.” (B) “The mouse is head-fixed while free to run on a wheel. An aluminum cone with 3D-printed wings makes direct contact with the edge of the headframe well to prevent the mouse’s tail from interfering with the probes and manipulators.” (C) “Image acquired after 6 probes have reached the surface of the brain.” (D) “Image acquired after the probes have reached their final depths.” (Allen Institute for Brain Science, 2019)

## Pre-processing Pipeline

To prepare the dataset for machine learning, a processing pipeline was implemented using the AllenSDK Python API, focusing on spike extraction, filtering, and storage across anatomical areas. The pipeline combined Allen's standard preprocessing (Allen Institute for Brain Science, 2019) with additional filtering, data structuring, and caching:

- **Data Access:** Sessions were loaded via the VisualBehaviorNeuropixelsProjectCache (v0.5.0), with metadata validation.
- **Unit and Channel Integration:** Units were merged by peak channel ID; only high-quality single units were retained ( $\text{SNR} > 1$ , Inter Spike Interval (ISI) (violations  $< 1\%$ , mean firing rate  $> 0.1 \text{ Hz}$ ).
- **Stimulus Extraction:** Change times were retrieved from stimulus tables; only active events were used.
- **Area-wise Structuring (HDF5):** Two-pass strategy: (1) count spikes per area to pre-allocate space, (2) stream spike times per unit into HDF5 groups (spikes, offsets, unit\_ids). This avoided large in-memory arrays.
- **Resource Management:** Memory was cleared between iterations; existing HDF5 outputs were skipped.
- **PSTH & Feature Extraction:** Spike times aligned to stimulus onset, binned at 25 ms over  $-1$  to  $+1.5$  s. Numba-optimized PSTHs were normalized and eight features extracted (six PSTH-based, two spike-train metrics) and stored as compressed .npz files.
- **Filtering & Caching:** Units with  $< 500$  samples were excluded; remaining regions reindexed, PSTHs, features, and one-hot labels cached. Sub-areas of the visual cortex, thalamus, and hippocampus were optionally isolated or concatenated. Due to computational and time constraints, midbrain regions were not included in the main analyses, and this study focused on cortex, thalamus, and hippocampal areas

## Alignment with Project Objective

The preprocessing strategy was designed to isolate electrophysiological features from anatomical assumptions. Structuring data per area and filtering by objective signal-quality metrics made the dataset suitable for testing whether each brain region exhibits a distinct electrophysiological signature. The resulting HDF5/npz architecture enables scalable, iterative analysis.

Due to the scale of recordings and large NWB file sizes (several gigabytes per session), some sessions could not be processed locally due to memory allocation failures (1.2–1.4 GiB per array on average), primarily in dense sessions exceeding 150 million spike timestamps. Eight of 90 sessions were excluded, while the remainder were successfully converted into area-wise HDF5 datasets containing tens of millions of spikes across 54 regions. These constraints highlight the trade-off between recording resolution and computational efficiency. Remaining areas and unit counts are listed in Table 1.

## Relevance to Project Objective

This preprocessing pipeline was designed to decouple functional neural signals from anatomical bias. By organizing per-area spike trains while excluding unstable or incomplete sessions, the dataset became suitable for training machine learning models that classify brain regions purely from electrophysiological signatures, testing the hypothesis that each anatomical area has a characteristic temporal firing structure.

## Exploratory data analysis

As observed in the correlation matrix (Fig. 2), electrophysiological activity exhibits both correlated and uncorrelated brain regions. Distinct brain areas contain neuron populations with differing firing dynamics and synaptic integration profiles (Tolosso et al., 2025; Kordecka et al., 2025; International Brain Laboratory, 2024). To quantify this diversity, eight spike-train features captured both firing rate and temporal structure (Table 2).

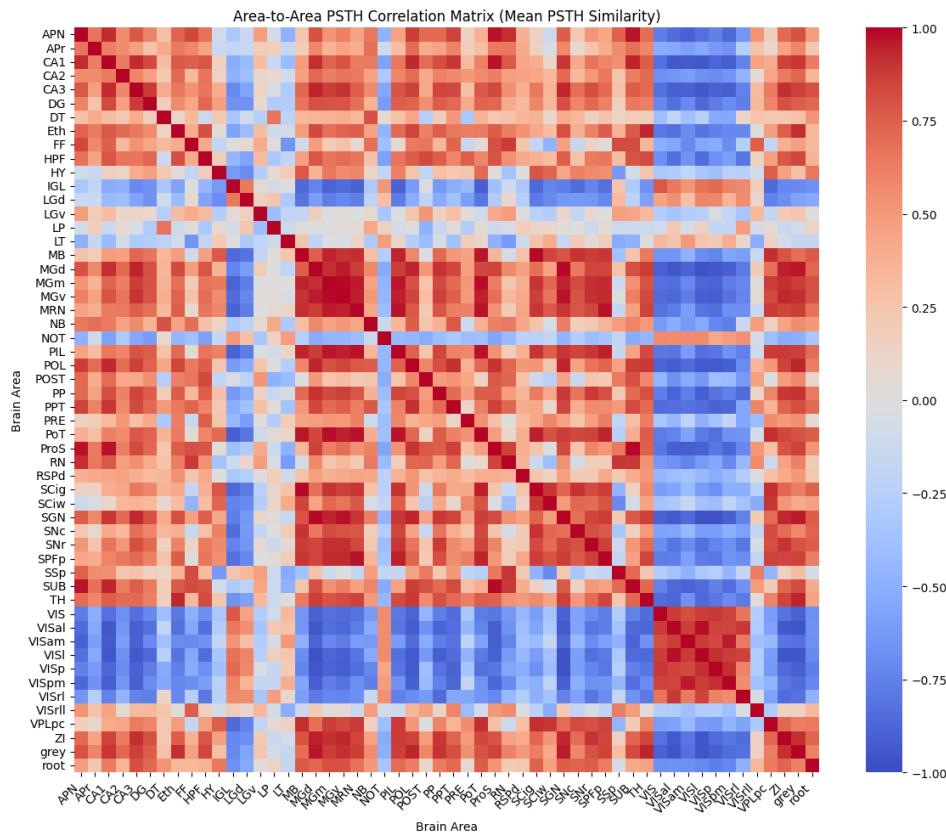
Visualization, including boxplots and inter-area dissimilarity matrices, is provided in the appendix. The omnibus dissimilarity matrix (Fig. 3) shows that ~52% of area pairs exhibited significantly distinct composite feature distributions (Bonferroni-corrected  $p < 3.49 \times 10^{-5}$ ; green), reflecting widespread divergence in neuronal response dynamics.

Differences between areas are visually reflected in PCA projections (Fig. 4, appendix). Distinct dynamics appear in the first four principal components alongside mean PSTHs per region (Fig. 5). Pairwise two-sample t-tests with FDR correction reveal region-specific dissimilarities: red clusters indicate similar response structure, green clusters significant divergence. Overall, pairwise dissimilarities accounted for ~42% of total variance per-area (CP#) relative to global principal components (PC#).

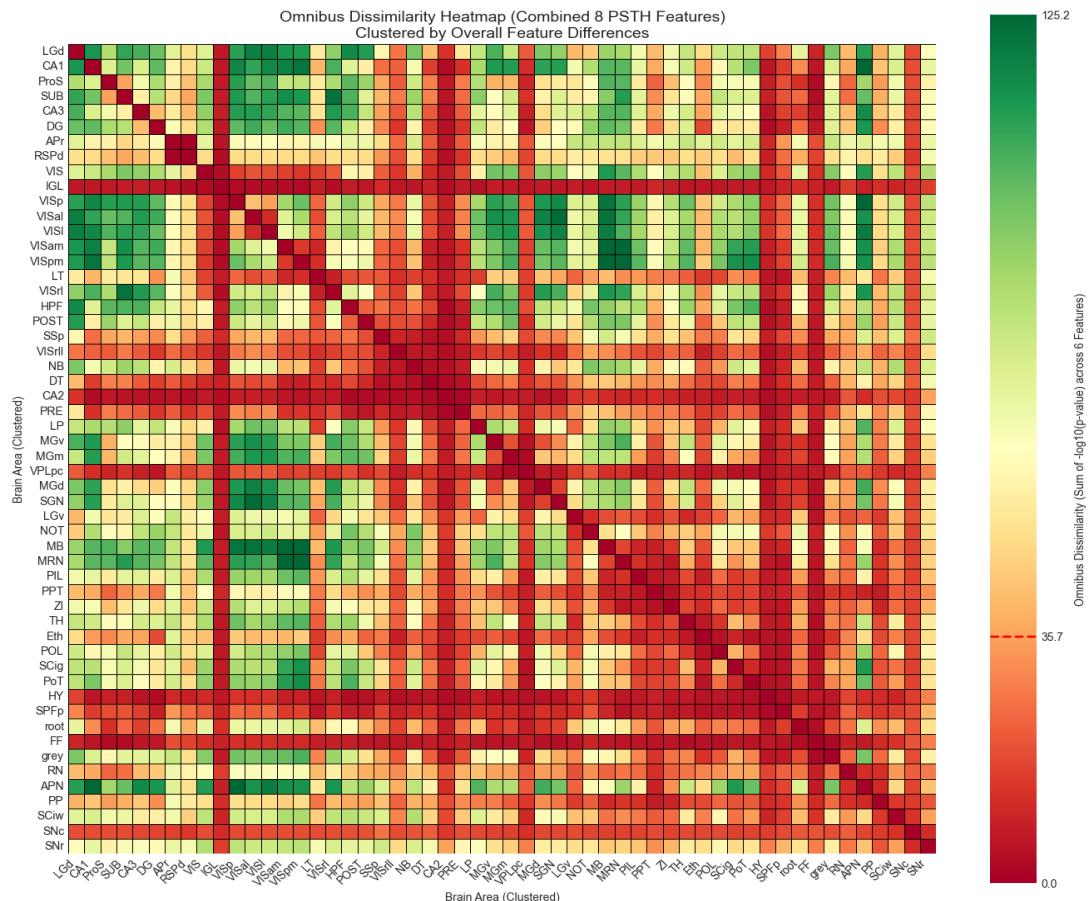
<b>Area</b>	<b>Units</b>	<b>Spikes</b>	<b>Superior Area</b>
APN	9,565	1,269,480,261	Midbrain
APr	5	223,780	Hippocampus
CA1	19,914	1,209,182,954	Hippocampus
CA2	8	411,786	Hippocampus
CA3	3,474	235,841,453	Hippocampus
DG	6,842	448,034,239	Hippocampus
DT	72	2,979,025	Thalamus
Eth	260	23,177,119	Thalamus
FF	31	2,458,316	Midbrain
HPF	1,204	82,053,056	Hippocampus
HY	22	1,661,485	Midbrain
IGL	5	365,049	Thalamus
LGd	1,926	194,829,984	Thalamus
LGv	413	46,457,949	Thalamus
LP	3,129	237,430,219	Thalamus
LT	164	11,408,785	Thalamus
MB	2,818	321,518,979	Midbrain
MGd	2,562	207,606,624	Thalamus
MGm	1,952	161,160,391	Thalamus
MGv	3,857	316,592,232	Thalamus
MRN	3,140	316,413,618	Midbrain
NB	863	50,051,197	Midbrain
NOT	792	86,779,031	Midbrain
PIL	1,024	108,733,042	Thalamus
POL	1,531	121,952,898	Thalamus
POST	1,608	120,278,708	Thalamus
PP	166	26,460,900	Thalamus
PPT	245	28,102,128	Midbrain
PRE	71	4,170,580	Hippocampus
PoT	1,676	139,070,738	Thalamus
ProS	2,745	238,939,645	Hippocampus
RN	161	30,851,455	Midbrain
RSPd	2	107,216	Hippocampus
SCig	1,613	125,873,608	Midbrain
SCIw	287	24,354,544	Midbrain
SGN	2,994	245,048,753	Thalamus
SNC	8	2,835,425	Midbrain
SNr	142	41,287,388	Midbrain
SPFp	88	7,193,848	Thalamus
SSp	153	9,124,051	Visual Cortex
SUB	3,995	339,655,571	Hippocampus

<b>Area</b>	<b>Units</b>	<b>Spikes</b>	<b>Superior Area</b>
TH	1,660	153,688,287	Thalamus
VIS	381	18,137,345	Visual Cortex
VISal	7,441	382,868,073	Visual Cortex
VISam	7,398	341,591,849	Visual Cortex
VISl	8,203	393,482,186	Visual Cortex
VISp	8,818	441,276,196	Visual Cortex
VISpm	8,885	372,855,455	Visual Cortex
VISrl	6,310	361,098,303	Visual Cortex
VISrll	64	4,015,008	Visual Cortex
VPLpc	31	2,344,822	Thalamus
ZI	525	57,638,270	Midbrain
grey	1,872	161,989,800	Midbrain
root	195	20,923,608	Midbrain

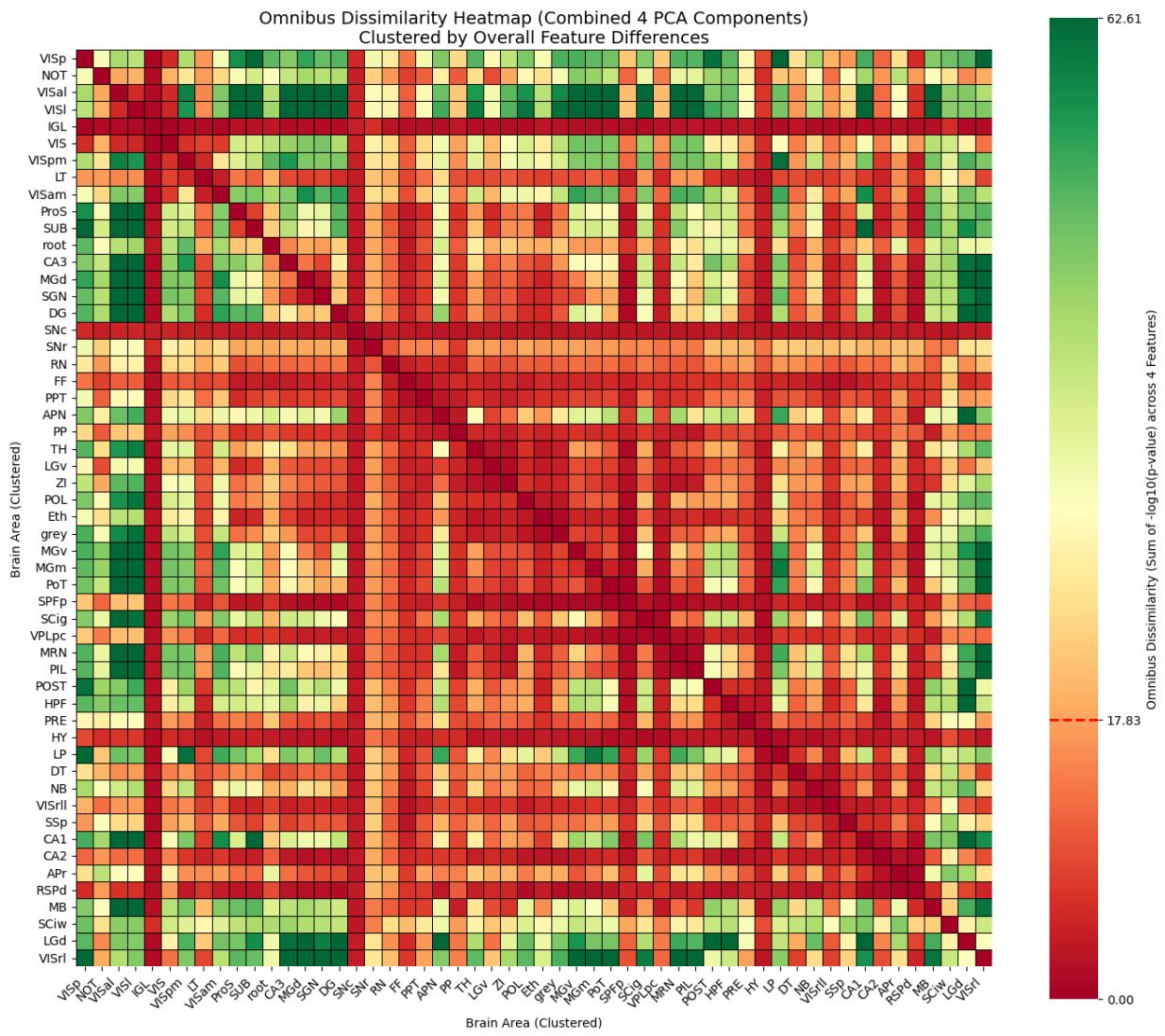
*Table 1, Unit count and spike count per area.*



*Figure 2: Correlation matrix, in red significant correlation and in blue uncorrelated*



*Figure 3: Omnibus classical features dissimilarity matrix, in red under 35,7 similar areas. In green above 35,7 dissimilar areas.*



*Figure 4: Omnibus first four principle components dissimilarity matrix, in red under 35,7 similar areas. In green above 35,7 dissimilar areas.*

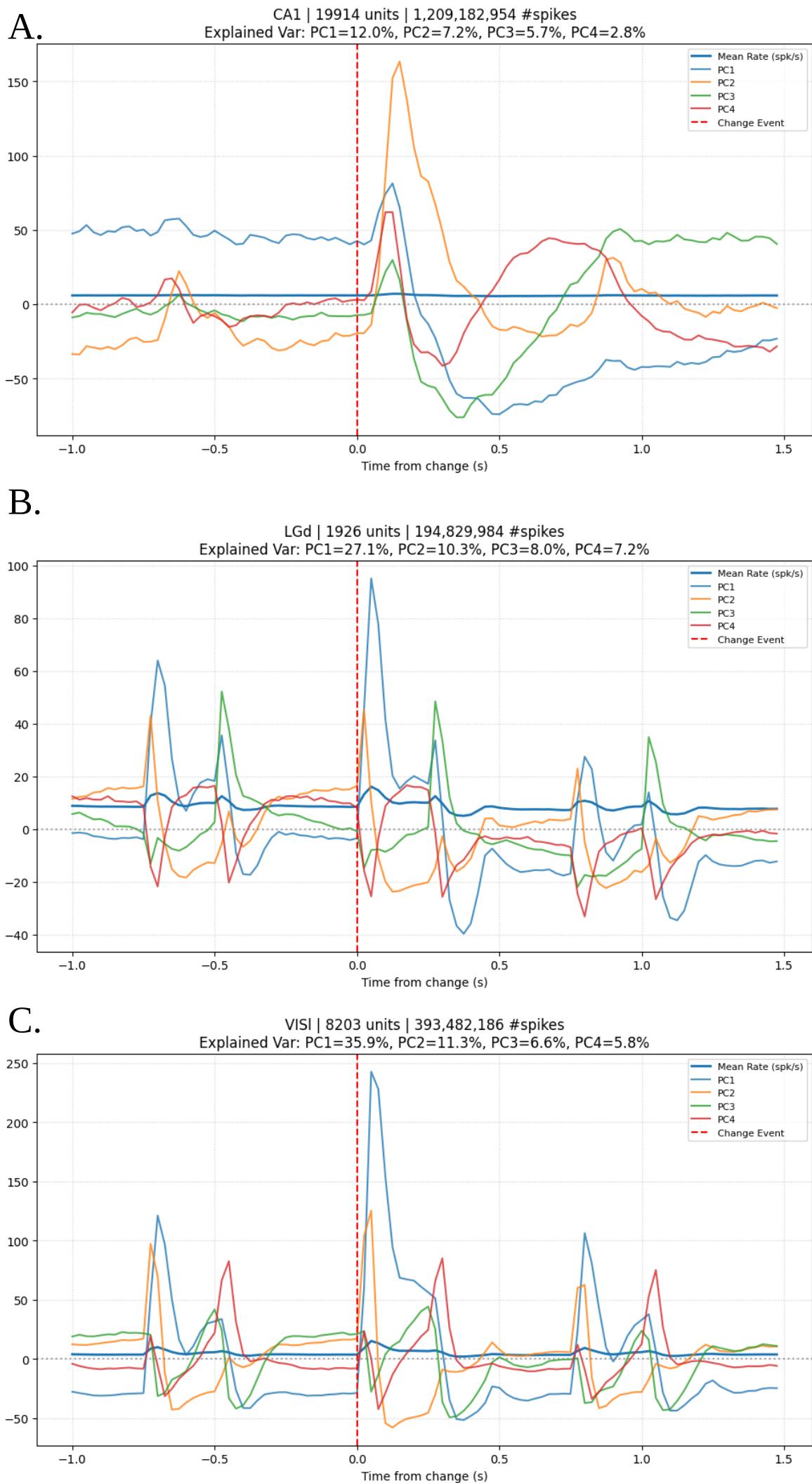


Figure 5: **A.** First four PC projections and PSTH from **A.** the ca1 sub area of the hippocampus, **B.** the lgd sub area of the thalamus and **C.** VISI sub area of the visual cortex

# Model Model structure, Training and testing methodologies.

## Input Representation

Neural activity was represented as PSTH's aligned to stimulus onset. Each trial spanned the window -1.0 s to 1.5 s with bin width 25 ms, resulting in 100 bins per trial. Eight scalar features were extracted to capture classical temporal response characteristics:

Feature	Description
$f_1$	<b>Timing of peak (s):</b> time of maximum firing rate relative to stimulus onset
$f_2$	<b>Peak firing rate (Z-score):</b> peak response normalized to baseline firing rate
$f_3$	<b>Response width:</b> duration during which firing rate remains above 50% of peak
$f_4$	<b>Response latency:</b> first time bin exceeding 3 standard deviations above baseline
$f_5$	<b>Response skewness:</b> asymmetry of post-stimulus firing rate distribution
$f_6$	<b>Sustained-to-transient ratio:</b> ratio of firing rate at 1.0 s post-stimulus to peak rate ( $r(1.0 s)/r_{peak}$ )
$f_7$	<b>Mean firing rate (spikes/s):</b> average spike rate over the response window
$f_8$	<b>CV ISI:</b> coefficient of variation of interspike intervals ( $\sigma_{ISI} / \mu_{ISI}$ ), measuring firing irregularity or spike timing variability

Table 2, 8 classical features

Thus all model variants share a two-stream architecture, with one stream modeling temporal structure in the PSTH sequence and the other processing scalar features. This design separates temporal signal processing from static feature analysis, minimizing cross-talk between data types. Represented as:

$$x = \{ X_{\text{PSTH}} \in \mathbb{R}^T, X_{\text{feat}} \in \mathbb{R}^8 \}.$$

## Loss and Optimization

The loss function incorporated class-weighted cross-entropy with a dampening coefficient ( $\alpha$ ) to stabilize gradients and reduce the influence of underrepresented classes:

$$L = -\sum_i \alpha * w_i^\beta * y_i * \log(\hat{y}_i)$$

where  $w_i^\beta$  denotes the class weights scaled by an exponent  $\beta$ , and  $\alpha \in [0.3, 2.0]$ ,  $\Delta\alpha = 0.05$  is a weight dampening factor tuned via Bayesian optimization.

Optimization used **Adam** with cosine learning rate decay:  $\eta_t = \eta_0 / 2 \cdot (1 + \cos(\pi t / T_{\max}))$ , and gradient clipping ( $\|g\|_2 \leq 1.0$ ).

## Multi headed streams and Hyper parameter optimization

By applying a Hyper Parameter Optimization (HPO) using Keras Bayesian Tuner(O’Malley, et al., 2019) separately and semi-uniformly across, Densely Multi-Layered Perceptron (DMLP; inspired by Rumelhart et al., 1986), a densely layered variant of Linear Discriminant Analysis (LDA; Fisher, 1936), a Convolutional Neural Network (CNN; LeCun et al., 1998), a Transformer model (Vaswani et al., 2017), and a Temporal Convolutional Network (TCN; Bai et al., 2018).

Data were split into 72/8/20 train-validation-test sets to preserve class balance across brain areas. Metrics included accuracy, balanced accuracy, and multi-class Area Under the Curve(AUC). Confusion matrices (row-normalized) provided insight into class-specific prediction reliability. Training and validation loss, accuracy, and AUC were monitored per epoch to assess convergence.

## Common Architecture for Two Streams:

- **Scalar feature branch:**  $X_{\text{feat}}$  (8 classical features) projected through a dense layer ( $S \in [16, 128]$ ) with ReLU activation, L2 regularization, dropout, and a learnable feature weighting coefficient before fusion.
- **Fusion and classification:** PTH/temporal representation concatenated with weighted scalar features, normalized, and fed into a dense classification head layers between 256-1024 units,  $\Delta=256$ , with ReLU, dropout, and a final softmax output. Denoted as:  
$$z = \{\text{NN}(X_{\text{PTH}}) \oplus \alpha \cdot \text{Dense}(X_{\text{feat}})\},$$
where  $\oplus$  indicates concatenation.

Optimization:

- Loss: class-weighted cross-entropy with dampening coefficient ( $\alpha \in [0.3, 2.0]$ )
- Optimizer: Adam with cosine learning rate decay and gradient clipping ( $\|g\|_2 \leq 1.0$ )
- Bayesian hyperparameter tuning: annealing dropout with an initial dropout probability  $p_0 \in [0.1, 0.7]$ ,  $\Delta p = 0.05$ , and gradually annealed toward a final value  $p_t = 0.05$ , L2 regularization ( $\lambda \in [10^{-6}, 1]$ ), feature weighting ( $\alpha \in [0.1, 10]$ ), learning rate  $\eta$  (log-uniform  $[10^{-7}, 1]$ ), batch size  $\in [16, 512]$ , total epochs  $\in [20, 150]$ , patience  $\in \{4, 36\}$  monitoring the validation Area Under the Curve(AUC) during HPO and validation loss during the HP optimized final training of the model, tuning trials  $\in [20, 100]$  and tuning epochs  $\in [3, 30]$ .

## Model Variants

### 1. DMLP

The DMLP baseline network integrates  $X_{\text{PSTH}}$  and  $X_{\text{feat}}$ . The PSTH branch received  $x_{\text{feat}}$  and passed them through 2 width-optimized dense layers ( $D1 \in [512, 4708]$ ,  $D2 \in [8, 512]$ ) both followed by ReLU, L2 regularization, batch normalization and annealed dropout. The fused representation was passed through the common classification head. This architecture captures nonlinear relationships between temporally averaged PSTHs and static population statistics without convolution or recurrent biases.

### 2. LDA DENSE

Replaces PSTH with D-dimensional LDA-transformed features ( $X_{\text{LDA}}$ ) to emphasize class-separating directions in neural population space. The LDA branch passed  $X_{\text{LDA}}$  through a fully connected block with layer widths optimized via Bayesian tuning ( $D1 \in [64, 512]$ ,  $\Delta D = 64$ ), followed by batch normalization, ReLU, and annealed dropout. The fused representation then followed the common classification head.

### 3. Transformer

The PSTH branch applied a linear embedding ( $E \in [32, 128]$ ) followed by positional encoding and  $B \in [1, 3]$  transformer blocks, each with multi-head self-attention ( $H \in [2, 8]$ ) and feedforward dimension  $F \in [128, 512]$ , ReLU activation, L2 regularization, and dropout. Sequence outputs were pooled via global average pooling to yield a fixed-length feature vector. This architecture captures long-range temporal dependencies

### 4. CNN

The PSTH branch applied a 1D convolutional stack with  $C \in [1, 3]$  layers, each with  $F \in [32, 128]$  filters, kernel size  $K \in \{5, 10, 20\}$ , ReLU activation, L2 regularization, batch normalization, and dropout. Max pooling ( $P \in \{2, 4\}$ ) reduced sequence length between layers, and sequence outputs were pooled via global average pooling to yield a fixed-length feature vector. This architecture captures local temporal patterns in PSTHs.

## 5. TCN

PSTH branch applied a Temporal Convolutional Network with  $L \in [2,8]$  residual levels, each containing dilated convolutions ( $F \in [16,128]$ ,  $K \in [3,7]$ ), ReLU, L2 regularization, and dropout. Dilation rates increase exponentially across residual levels ( $2^0, 2^1, \dots, 2^{L-1}$ ). Outputs were pooled, fused with scalar features, and passed to the classification head. This architecture captures long-range temporal dependencies via dilated convolutions.

# Results

The performance of classical and deep learning models for brain region classification was evaluated (Table 3), focusing on weighted accuracy to account for imbalanced region representation. Classical classifiers (SGD, Passive Aggressive, Perceptron, Nearest Centroid, MultinomialNB) achieved low weighted accuracy (0.03-0.19), indicating limited generalization across sparsely sampled regions.

Deep learning models substantially improved performance. DLMP reached weighted accuracy of 0.11 overall, increasing in hippocampus (0.24) and thalamus (0.33). LDA Dense achieved 0.21 overall and 0.51 in thalamus. Transformers showed some generalization (weighted accuracy 0.35), particularly in thalamus (0.56) and visual areas (0.39). CNNs performed relatively well in thalamus (0.46) and hippocampus (0.27), while TCN captured temporal some firing patterns (0.12 overall; 0.31 in thalamus). For context, the Dense model from Tolosso et al. (2025) achieved weighted accuracies of 0.44, 0.26, and 0.25 on the same dataset for visual, hippocampal, and thalamic units, respectively.

Training and validation loss differed by ~10% on average (figure 7), suggesting mild overfitting for several models, particularly those with higher parameter counts. As seen by the divergence of the training and validation accuracy from TCN, MLP and CNN (figure 7).

The HPO search space was visualized using a custom PyQt-based HPO Explorer (Figure 8), which displays trial-wise metrics including training and validation AUC, loss, and accuracy. This visualization facilitates detection of promising hyperparameter regions. The characteristic oscillations in loss and AUC, combined with a gradual upward trend in accuracy, reflect the optimizer's exploration of both effective and ineffective hyperparameter configurations. The best hyper parameters per model and area can be found in the appendix.

Normalized confusion matrices highlight class-specific performance. The LDA Dense model achieved moderate recall across most regions, with confusion between adjacent cortical areas (Figure 6.A). Transformers captured the thalamic and hippocampal activity moderately well but showed confusion among sensory cortical units (Figure 6.B; Appendix, Confusion matrices and training plots). CNNs and DMLP displayed more distributed misclassification patterns, reflecting their focus on local temporal dependencies (Figure 6.B, 6.E and appendix, Confusion matrices and training plots). The TCN achieved a comparatively balanced classification across regions, capturing

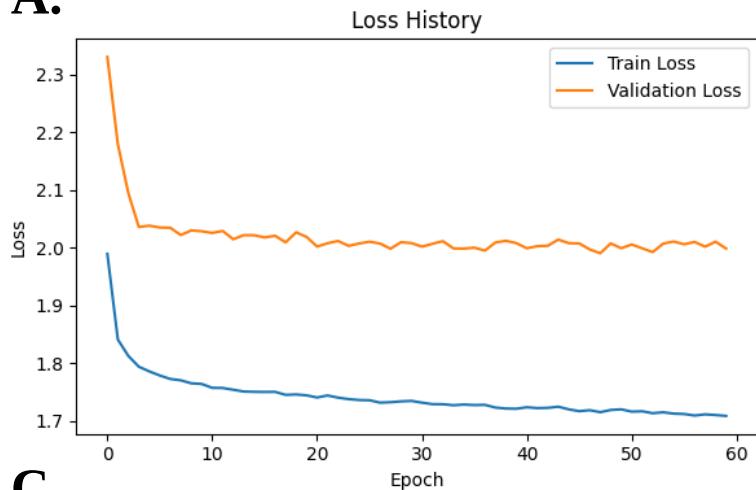
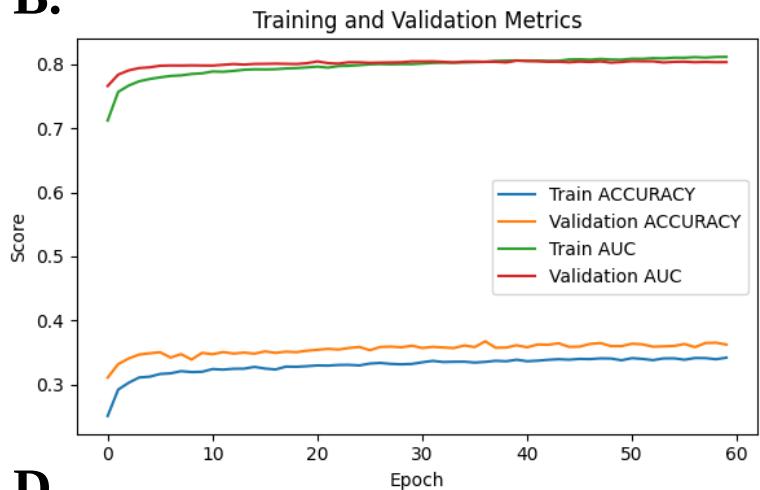
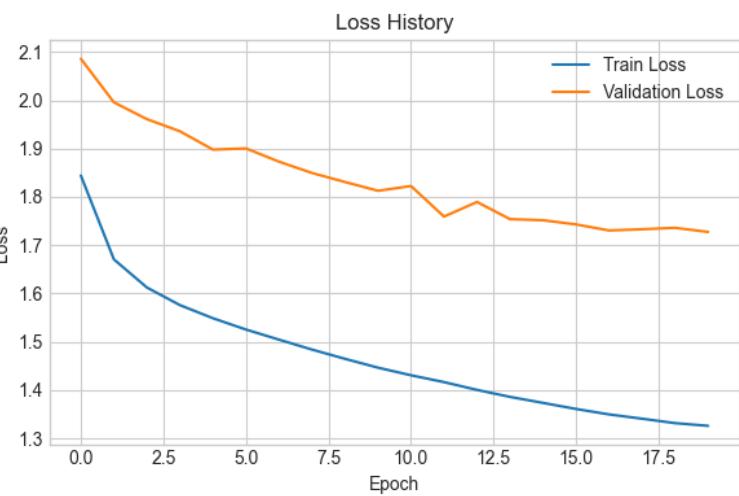
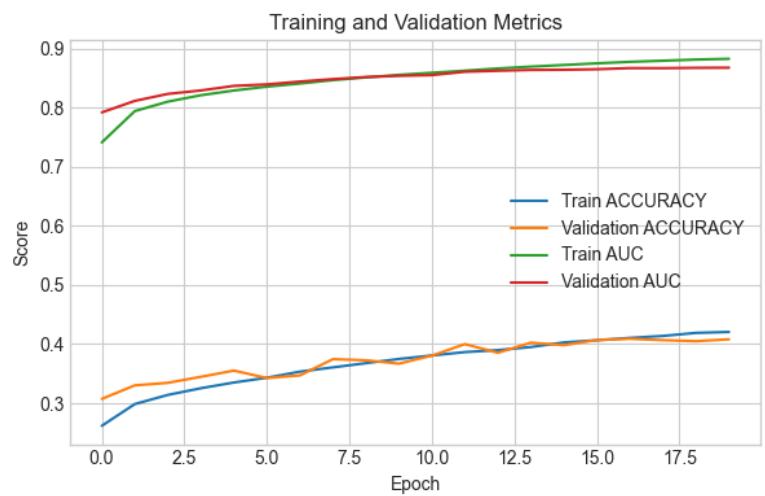
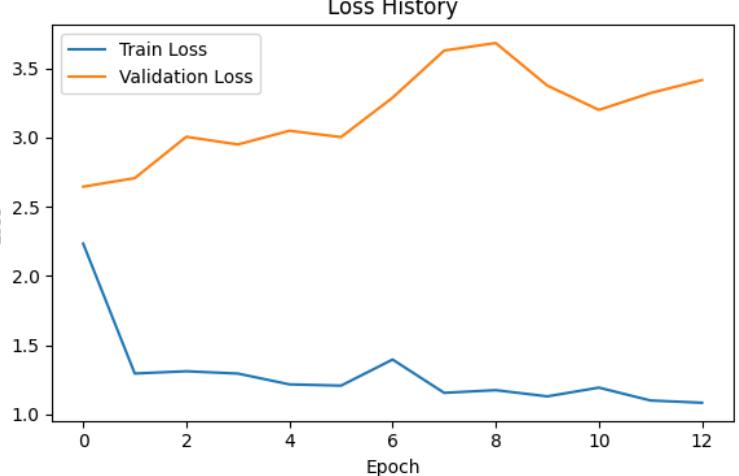
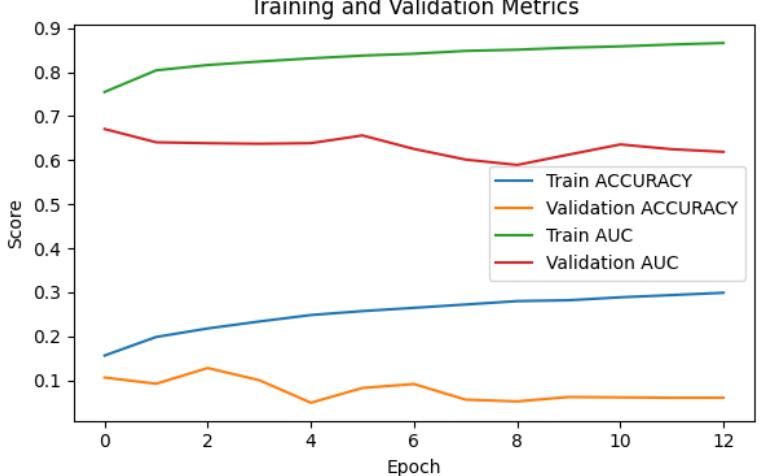
long-range temporal dependencies while maintaining some discrimination between neighboring cortical and subcortical areas (Figure 6.C).

Overall, deep learning models, particularly Transformers and LDA Dense, leveraged electrophysiological patterns somewhat effectively for brain region classification, outperforming classical approaches while revealing region-specific some discriminability and some susceptibility to overfitting.

<b>Algorithm</b>	<b>Region</b>	<b>Accuracy</b>	<b>Weighted Accuracy</b>	<b>AUC</b>
<b>SGD</b>	All units	—	0.08	—
<b>Passive Aggressive</b>	All units	—	0.09	—
<b>Perceptron</b>	All units	—	0.10	—
<b>Nearest Centroid</b>	All units	—	0.03	—
<b>MultinomialNB</b>	All units	—	0.19	—
<b>DLMP</b>	All units	0.09	0.11	0.65
	Visual	0.19	0.18	0.55
	Hippocampus	0.40	0.24	0.61
	Thalamus	0.43	0.33	0.67
<b>LDA Dense</b>	All units	0.35	0.21	0.80
	Visual	0.32	0.32	0.69
	Hippocampus	0.52	0.31	0.66
	Thalamus	0.51	0.51	0.76
<b>Transformer</b>	All	0.39	0.35	0.86
	Visual	0.38	0.39	0.75
	Hippocampus	0.30	0.43	0.72
	Thalamus	0.55	0.56	0.80
<b>CNN</b>	All	0.19	0.26	0.80
	Visual	0.25	0.24	0.63
	Hippocampus	0.54	0.27	0.67
	Thalamus	0.45	0.46	0.72
<b>TCN</b>	All	0.10	0.12	0.67
	Visual	0.17	0.18	0.58
	Hippocampus	0.15	0.26	0.63
	Thalamus	0.26	0.31	0.66
<b>Dense</b> (Tolosso et al, 2025)	Visual	-	0.44	-
	Hippocampus	-	0.26	-
	Thalamus	-	0.25	-

*Table 3, Performance comparison of classical and deep learning models in classifying neural units by brain region. Reported metrics include overall accuracy, weighted accuracy, and area under the*



**A.****B.****C.****D.****E.****F.**

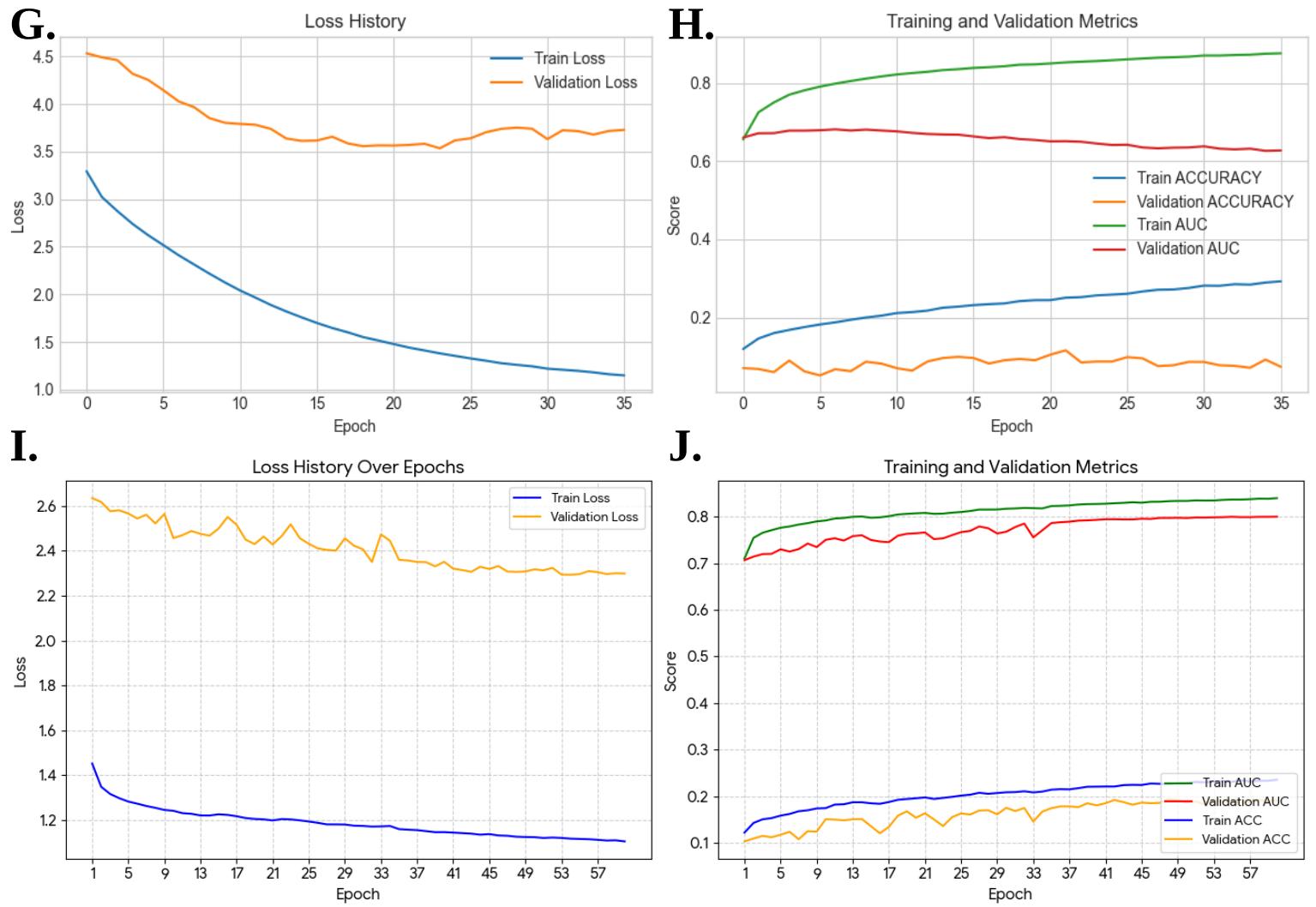
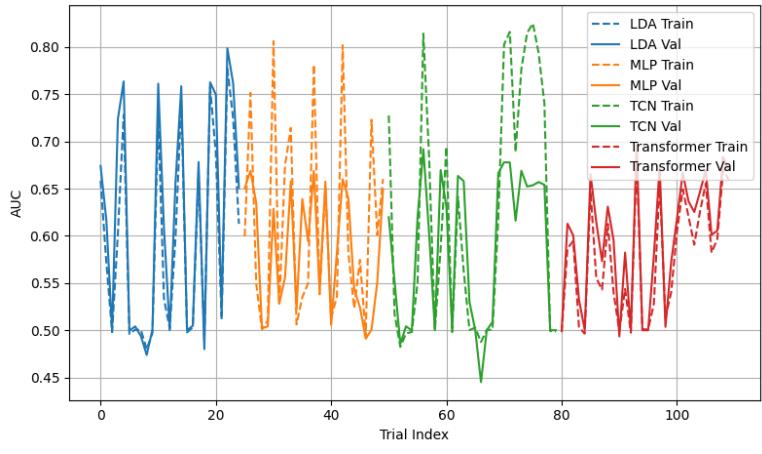


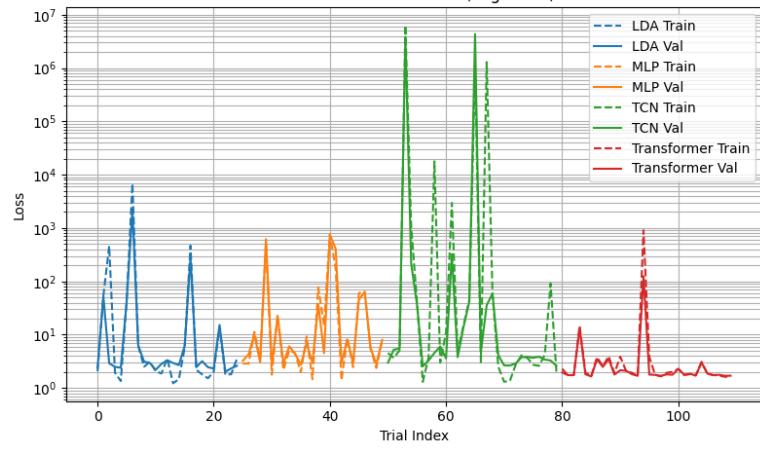
Figure 7: (A, C, E, G, I) Training histories and (B, D, F, H, J) corresponding training and validation metrics for LDA Dense, Transformer, TCN, Dense MLP, and CNN models, respectively.

**A.**

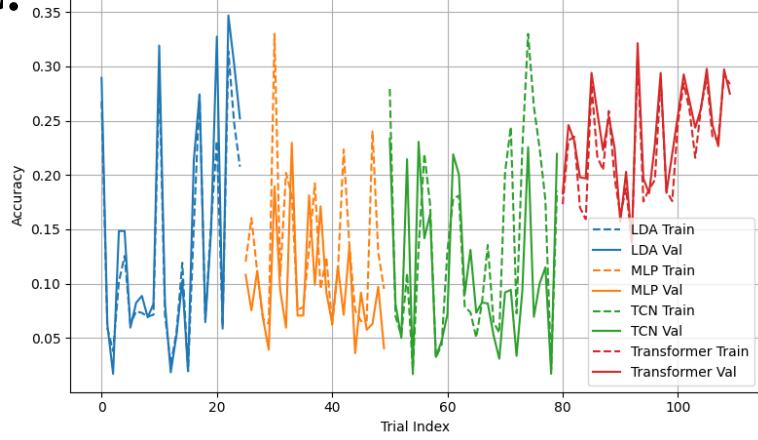
Train vs Validation AUC

**B.**

Train vs Validation Loss (Log Scale)

**C.**

Train vs Validation Accuracy



*Figure 8, Bayesian HPO trial iterations for all neural network models are shown sequentially, depicting (A) training and validation AUC, (B) training and validation loss on a logarithmic scale, and (C) training and validation accuracy.*

## Conclusion

This study addressed probe localization in large-scale electrophysiology, where current methods based on post-hoc histology and registration to the Allen CCFv3 atlas are limited by observer bias and individual biological variability. The central objective was to develop and evaluate a data-driven framework capable of identifying brain regions based solely on their intrinsic electrophysiological signatures, bypassing traditional anatomical mapping.

To test this hypothesis, we utilized the Allen Brain Observatory Neuropixels dataset, processing spike-train data to extract both classical temporal features (e.g., peak latency, firing rate, CV\_ISI) and full peristimulus time histograms (PSTHs). I then systematically compared the performance of classical classifiers against a suite of hyperparameter-optimized, two-stream deep learning models (DMLP, LDA Dense, CNN, TCN, and a Transformer).

The results provide a clear proof-of-concept for this histology-independent approach. We found that:

- Classical classifiers (e.g., SGD, MultinomialNB) failed to generalize, with weighted accuracies below 0.19, demonstrating the task's complexity.
- Deep learning models offered substantial improvements, confirming that unique, classifiable information is embedded in the neural firing patterns.
- The Transformer model achieved the highest overall performance, with a weighted accuracy of 0.35 and an AUC of 0.86. The LDA Dense model also showed strong performance (weighted accuracy 0.21, AUC 0.80).
- Performance was region-dependent, with models showing the highest discriminability in the thalamus (0.56 weighted accuracy for the Transformer), suggesting its response profile is particularly distinct.
- Compared to the Dense model benchmark reported by Tolosso et al. (2025), which achieved weighted accuracies of 0.25–0.44 across regions, the two-stream architectures provide a competitive advantage by integrating complementary temporal and scalar neural features. This separation of streams might have contributed to improved region-specific discriminability, particularly in thalamic and hippocampal populations.

These findings leans towards validation of the study's core assumption that anatomical regions possess quantifiable electrophysiological signatures. While the overall accuracy indicates that the problem is far from solved, the success of models like the Transformer demonstrates the potential of using deep learning to capture complex temporal dependencies as a basis for anatomical classification.

## Limitations and Future Directions

The overall accuracy, while well above chance, is modest, and confusion matrices indicate that models struggled to separate adjacent or functionally similar sub-regions. Mild overfitting was observed in several models, and computational hardware limitations prevented the processing of the complete dataset, pointing to practical scaling issues.

Future research should focus on improving model separability. This could involve:

- **Richer Input Features:** Incorporating Local Field Potential (LFP) data alongside spike trains.
- **Advanced Architectures:** Exploring graph neural networks (GNNs) to model the spatial relationships between adjacent probe channels.
- **Cross-Validation:** Testing the trained models on datasets from different laboratories or using different stimulus protocols to assess their robustness and generalizability.
- **Expanded Brain Coverage:** Incorporating midbrain recordings. Future work could extend the framework to these regions to evaluate whether electrophysiological signatures generalize across major brain structures.

This approach offers a path toward a more objective, scalable, and reproducible method for localizing neural recordings, which could ultimately serve as a vital quality-control tool to complement traditional histological methods. Beyond methodological advances, it has the ethical advantage of potentially reducing the need for post-mortem procedures, thereby limiting animal suffering in neuroscience research.

## References

Allen Institute for Brain Science. (2019) *Allen Brain Observatory – Neuropixels Visual Coding: Technical White Paper, Version 1.0 (October 1, 2019)*. Available at:

<https://portal.brain-map.org/explore/circuits/visual-coding-neuropixels>

Bai, S., Kolter, J.Z. and Koltun, V., 2018. *An empirical evaluation of generic convolutional and recurrent networks for sequence modeling*. arXiv preprint arXiv:1803.01271.

Fisher, R.A. (1936) The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, 7, 179-188. <http://dx.doi.org/10.1111/j.1469-1809.1936.tb02137.x>

Tolossa, G. B., Nowak, L. M., Kaneko, K., Iyer, R., Pachitariu, M. and Harris, K. D. (2025) ‘Neurons throughout the brain embed robust signatures of anatomical location’, *eLife*, 14, e101506. doi:10.7554/eLife.101506.

Kordecka, K., Górska, U., Krauz, R. and Godlewski, M. (2025) ‘Comparison of physiological brain responses evoked by electrical and visual stimulation’, *Frontiers in Neuroscience*, 19, Article 12054683. [doi:10.3389/fnins.2025.12054683](https://doi.org/10.3389/fnins.2025.12054683).

Kiernan, J.A. (2015). Histological and Histochemical Methods, Theory and Practice, 5th Ed. Banbury: Scion Publishing Ltd.

International Brain Laboratory , Banga Kush, Benson Julius, Bhagat Jai, Biderman Dan, Birman Daniel, Bonacchi Niccolò, Bruijns Sebastian A, Buchanan Kelly, Campbell Robert AA, Carandini Matteo, Chapuis Gaëlle A, Churchland Anne K, Davatolagh M Felicia, Lee Hyun Dong, Faulkner Mayo, Gercek Berk, Hu Fei, Huntenburg Julia, Hurwitz Cole, Khanal Anup, Krasniak Christopher, Langfield Christopher, Meijer Guido T, Miska Nathaniel J, Mohammadi Zeinab, Noel Jean-Paul, Paninski Liam, Pan-Vazquez Alejandro, Roth Noam, Schartner Michael, Socha Karolina, Steinmetz Nicholas A, Svoboda Karel, Taheri Marsa, Urai Anne E, Wells Miles, West Steven J, Whiteway

Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.

Matthew R, Winter Olivier, Witten Ilana B (2024) Reproducibility of in vivo electrophysiological measurements in mice *eLife* 13:RP100840 <https://doi.org/10.7554/eLife.100840.1>

O’Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L. et al. (2019) *KerasTuner*. Available at: <https://github.com/keras-team/keras-tuner> accessed: 2025

Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986. *Learning representations by back-propagating errors*. Nature, 323(6088), pp.533–536.

Steinmetz NA, Aydin C, Lebedeva A, Okun M, Pachitariu M, Bauza M, Beau M, Bhagat J, Böhm C, Broux M, Chen S, Colonell J, Gardner RJ, Karsh B, Kloosterman F, Kostadinov D, Mora-Lopez C, O’Callaghan J, Park J, Putzeys J, Sauerbrei B, van Daal RJJ, Volland AZ, Wang S, Welkenhuysen M, Ye Z, Dudman JT, Dutta B, Hantman AW, Harris KD, Lee AK, Moser EI, O’Keefe J, Renart A,

Svoboda K, Häusser M, Haesler S, Carandini M, Harris TD. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science*. 2021 Apr 16;372(6539):eabf4588. doi: 10.1126/science.abf4588. PMID: 33859006; PMCID: PMC8244810.

Tolossa GB, Schneider AM, Dyer E, Hengen KB. Neurons throughout the brain embed robust signatures of their anatomical location into spike trains. *Elife*. 2025 Jun 27;13:RP101506. doi: 10.7554/eLife.101506. PMID: 40577195; PMCID: PMC12204685.

Tsien, J.Z., 2016. *Cre-Lox neurogenetics: 20 years of versatile applications in brain research and counting...* *Frontiers in Genetics*, 7, p.19. Available at: <https://doi.org/10.3389/fgene.2016.00019>

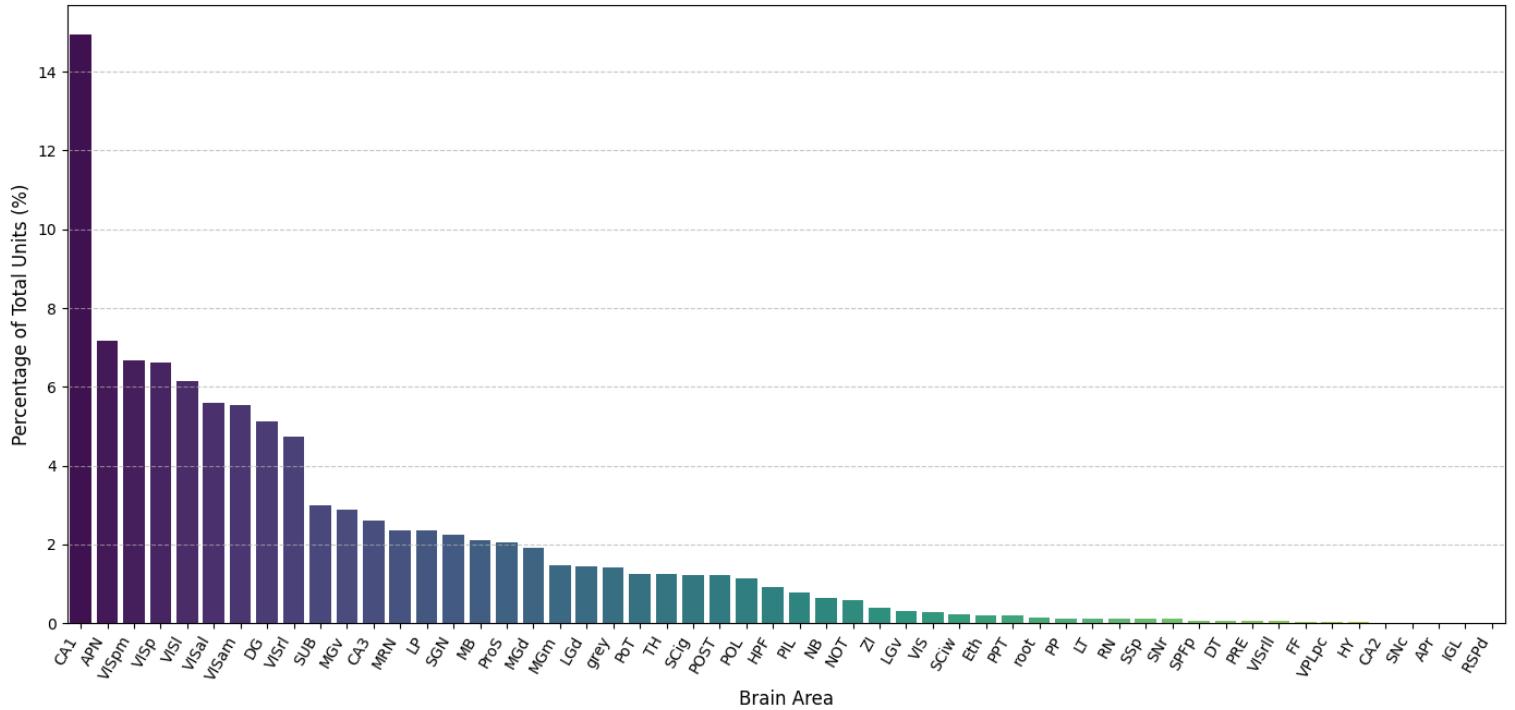
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. *Attention is all you need*. arXiv preprint arXiv:1706.03762. Available at: <https://doi.org/10.48550/arXiv.1706.03762>

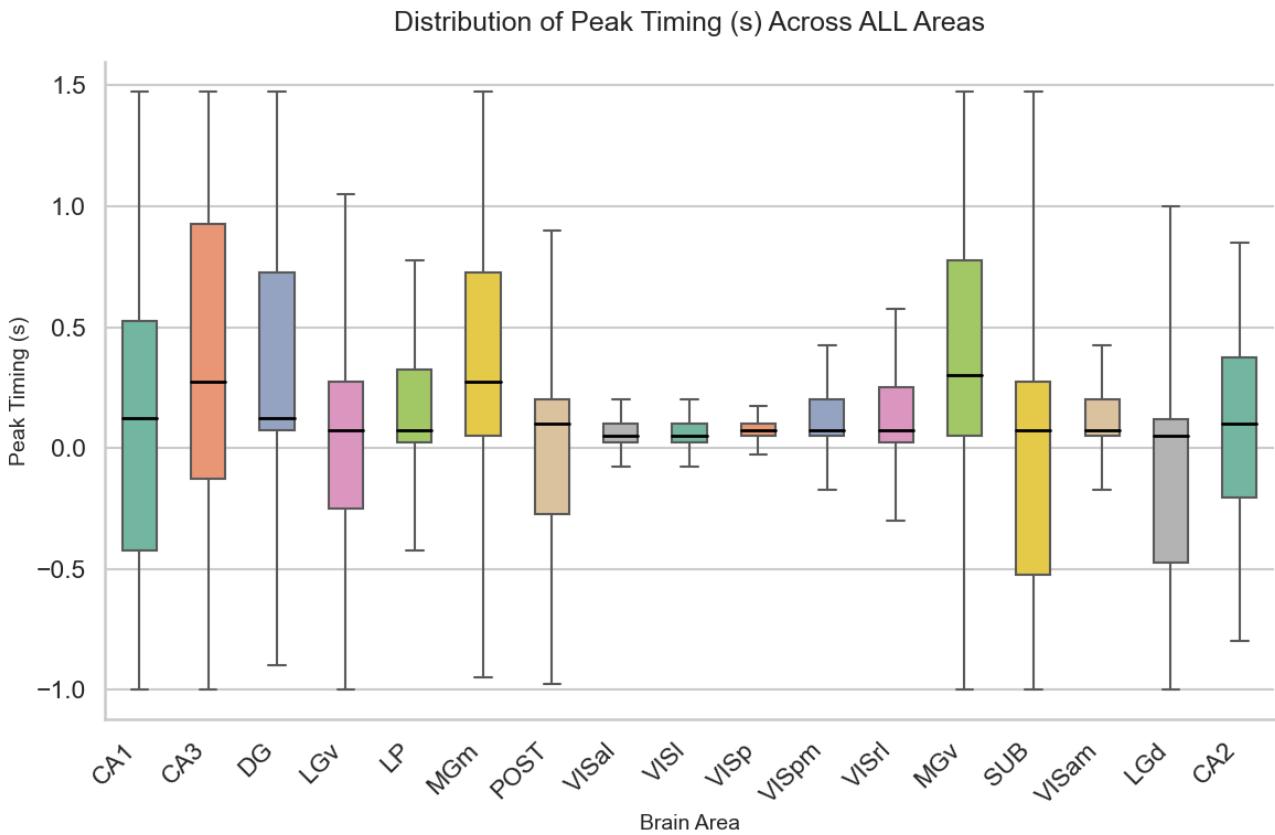
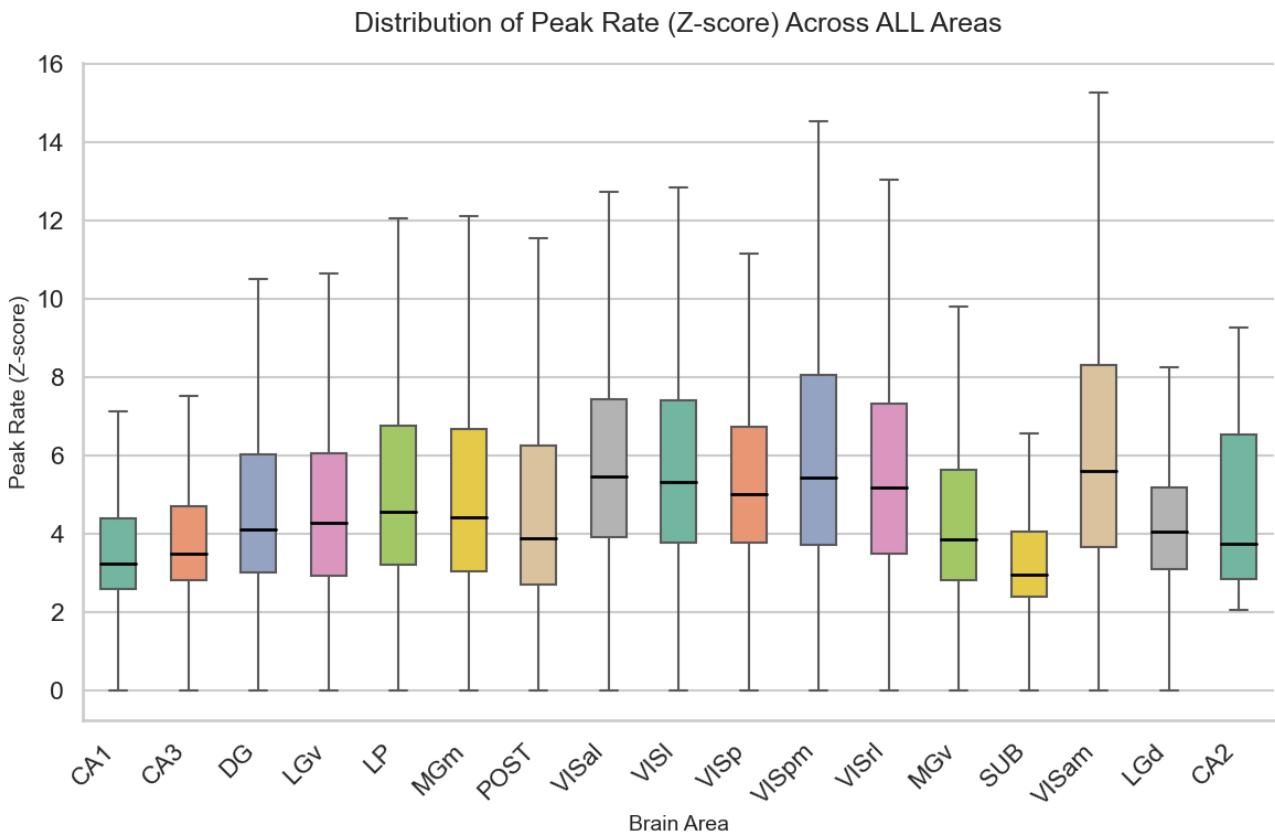
Wang Q, Ding SL, Li Y, Royall J, Feng D, Lesnar P, Graddis N, Naeemi M, Facer B, Ho A, Dolbeare T, Blanchard B, Dee N, Wakeman W, Hirokawa KE, Szafer A, Sunkin SM, Oh SW, Bernard A, Phillips JW, Hawrylycz M, Koch C, Zeng H, Harris JA, Ng L. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*. 2020 May 14;181(4):936-953.e20. doi: 10.1016/j.cell.2020.04.007. Epub 2020 May 7. PMID: 32386544; PMCID: PMC8152789.

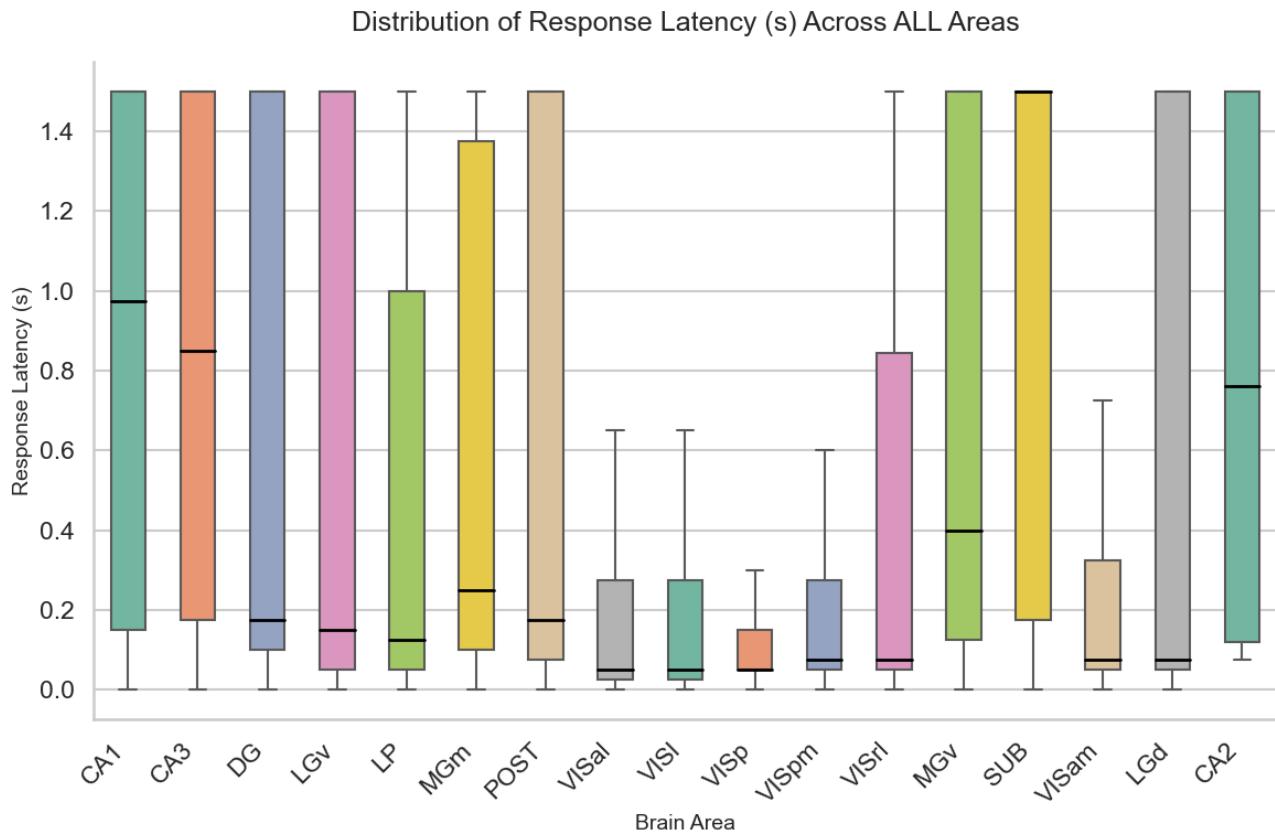
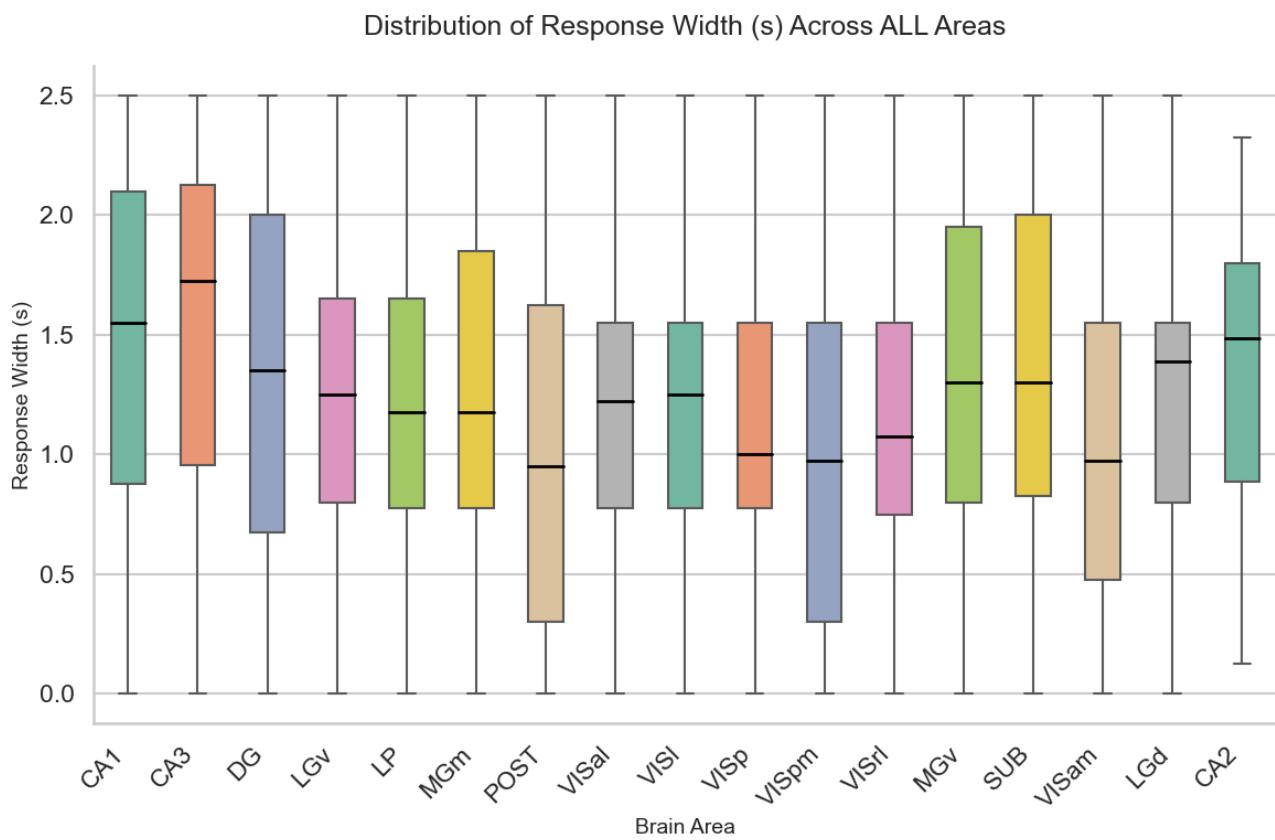
# Appendix

## Unit distributions and feature boxplots

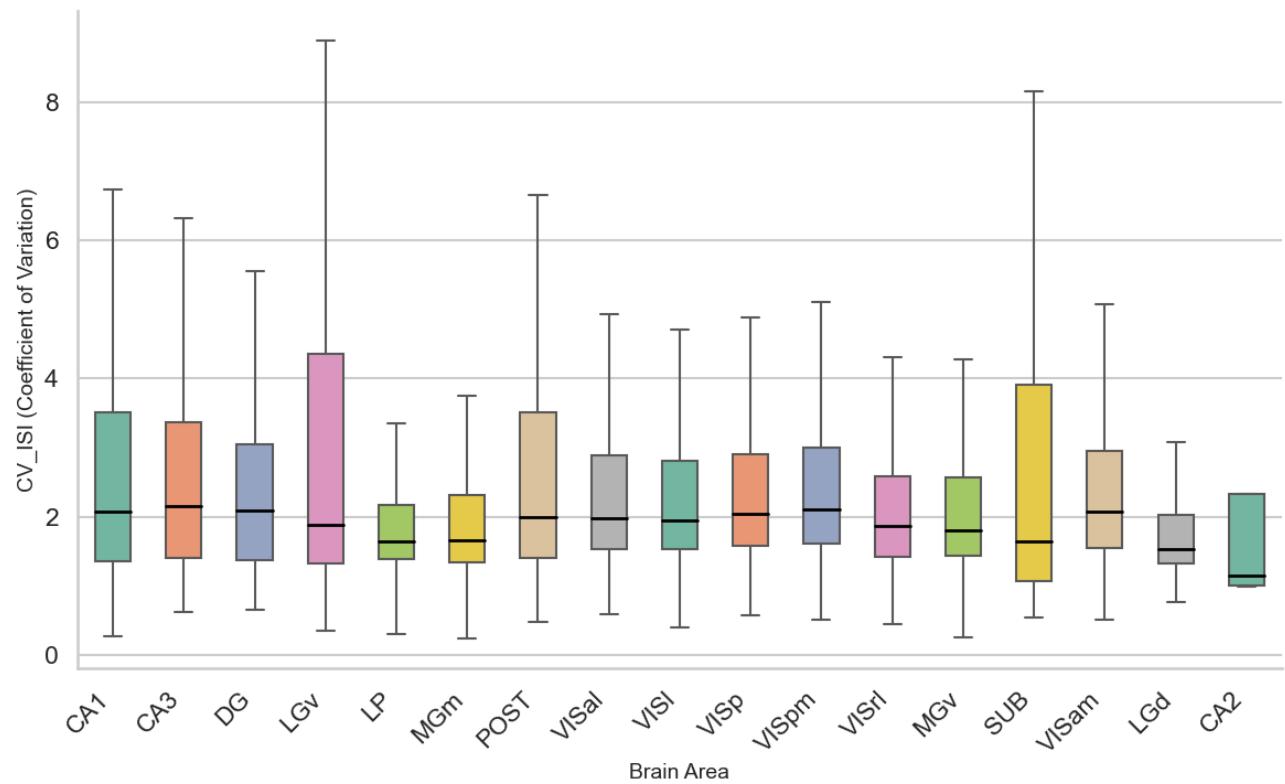
Unit Distribution Across Brain Areas (Total Units: 133310)



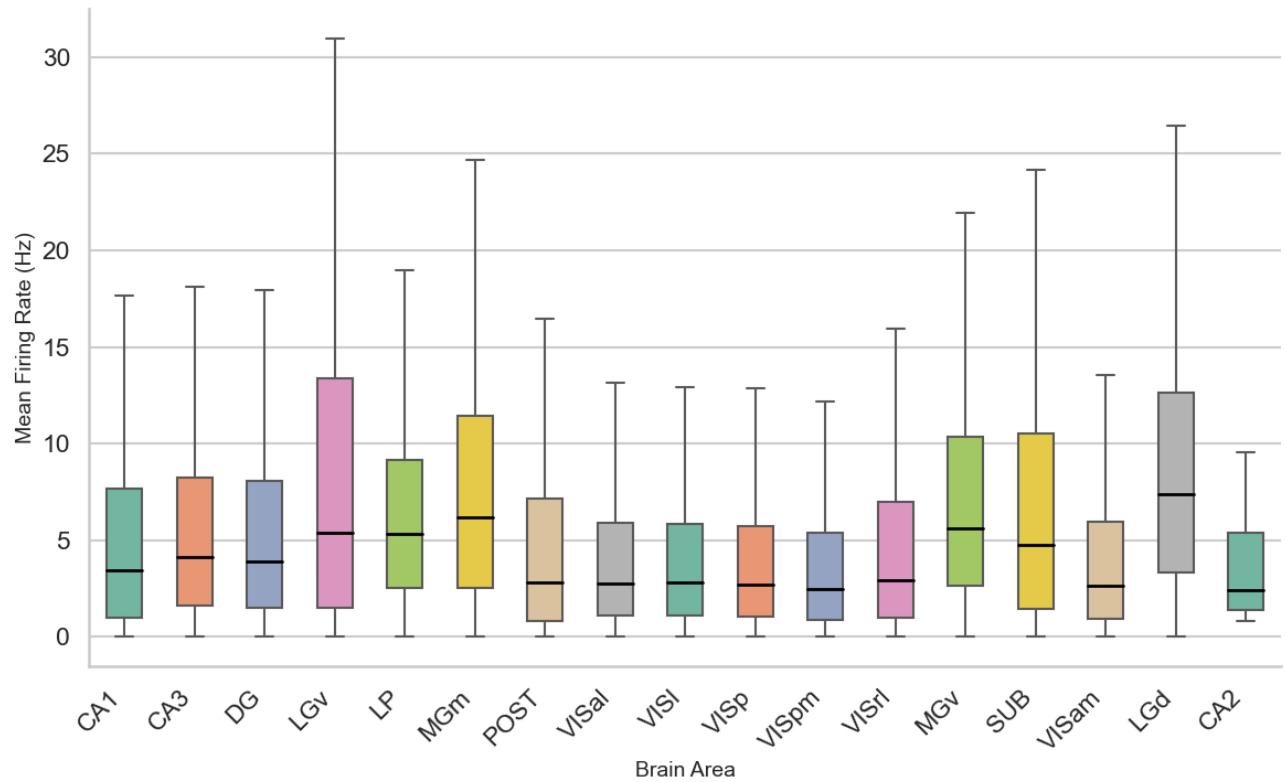




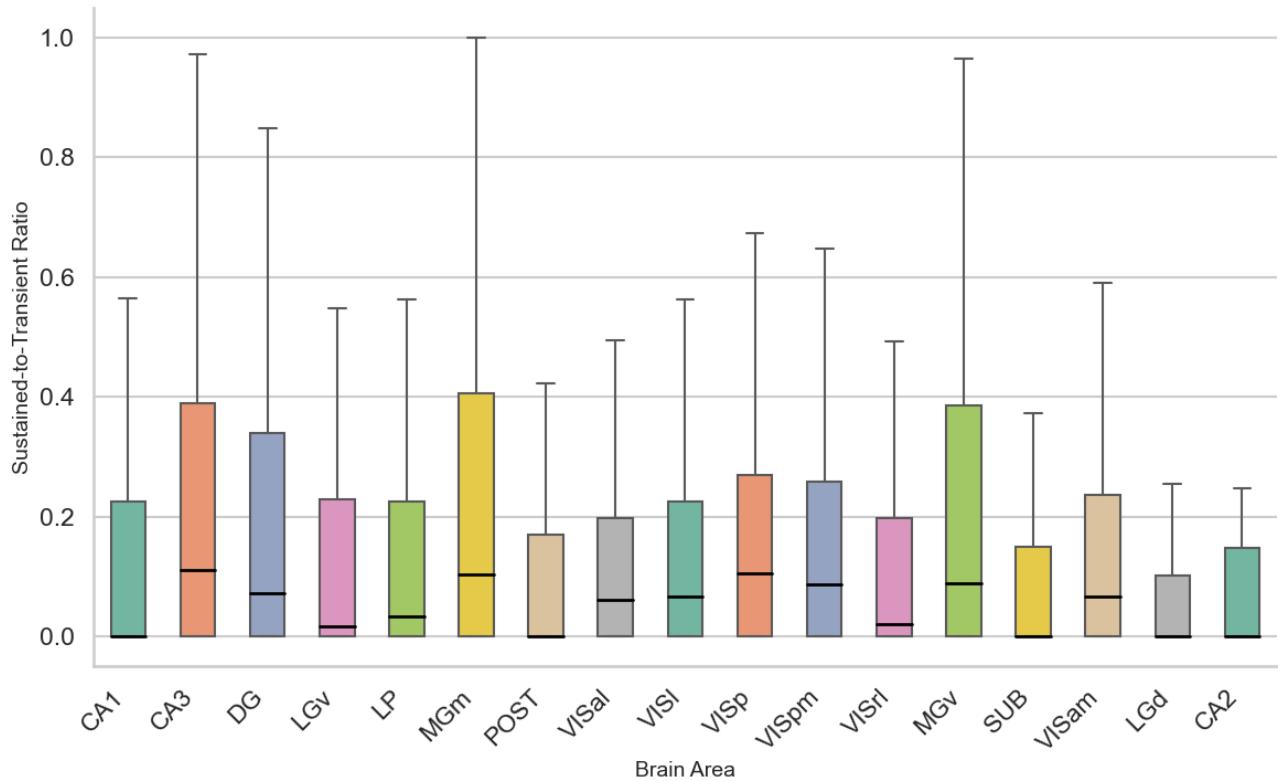
Distribution of CV\_ISI (Coefficient of Variation) Across ALL Areas



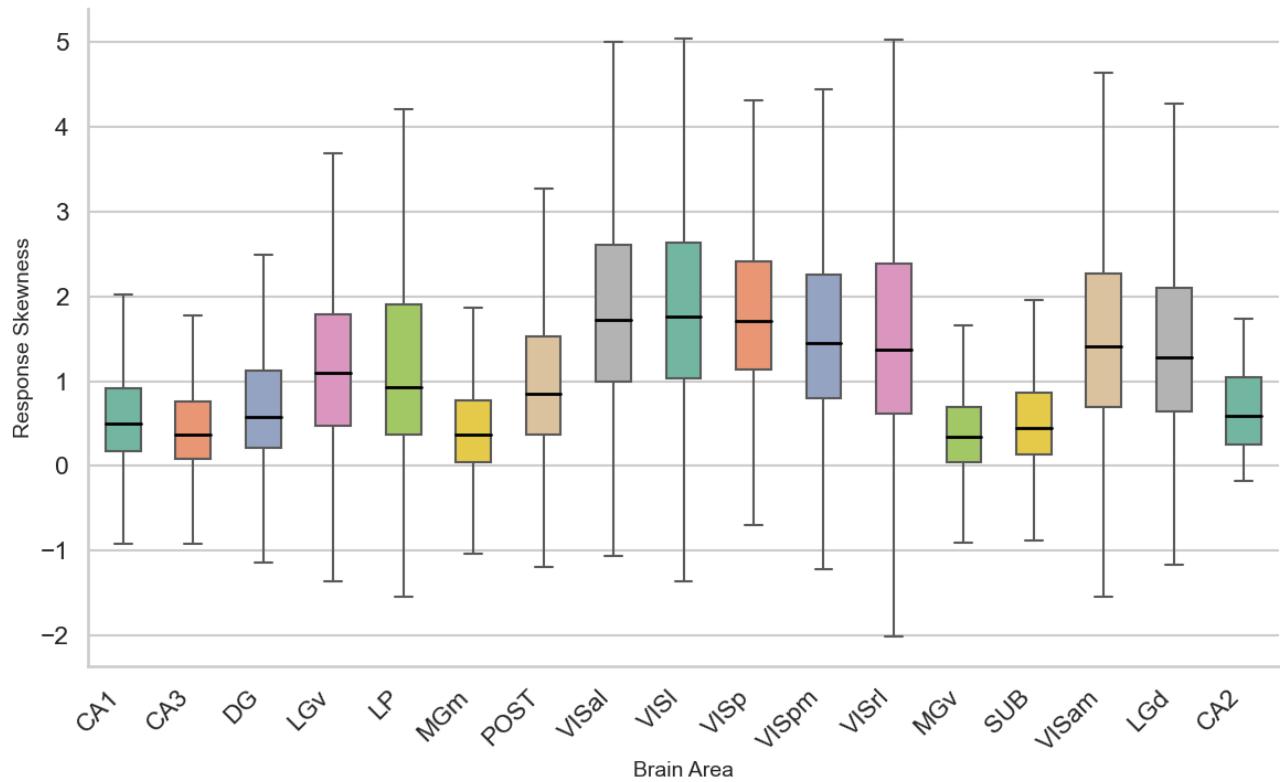
Distribution of Mean Firing Rate (Hz) Across ALL Areas



Distribution of Sustained-to-Transient Ratio Across ALL Areas

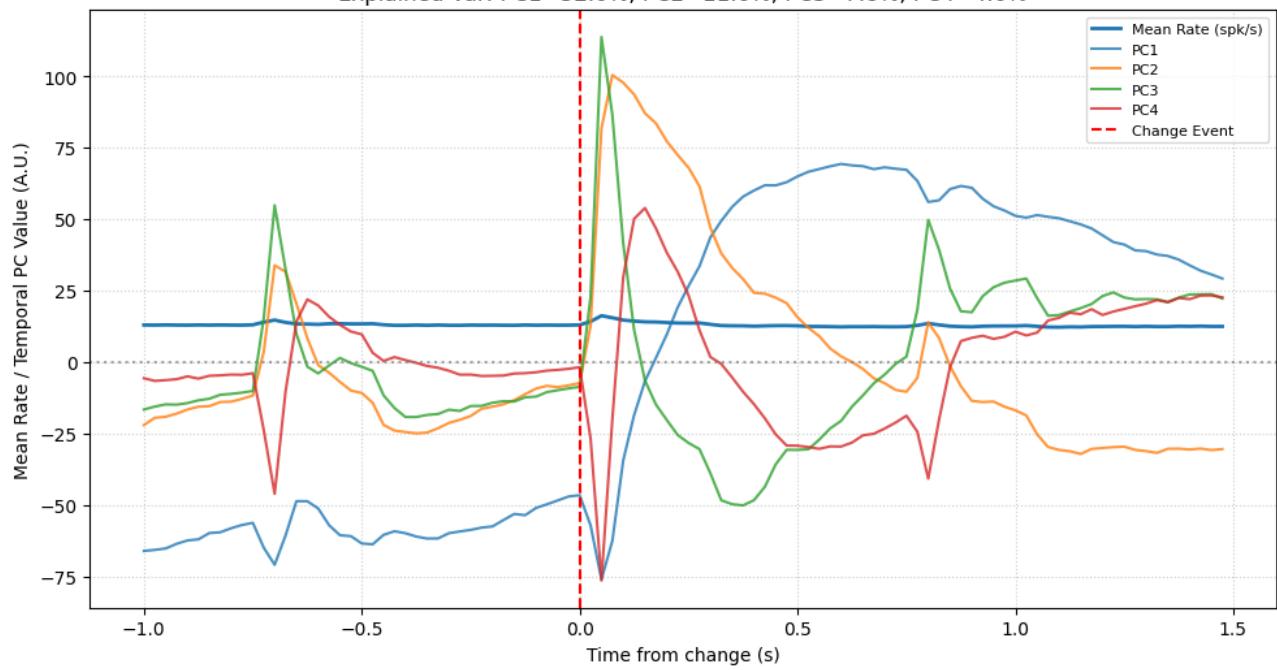


Distribution of Response Skewness Across ALL Areas

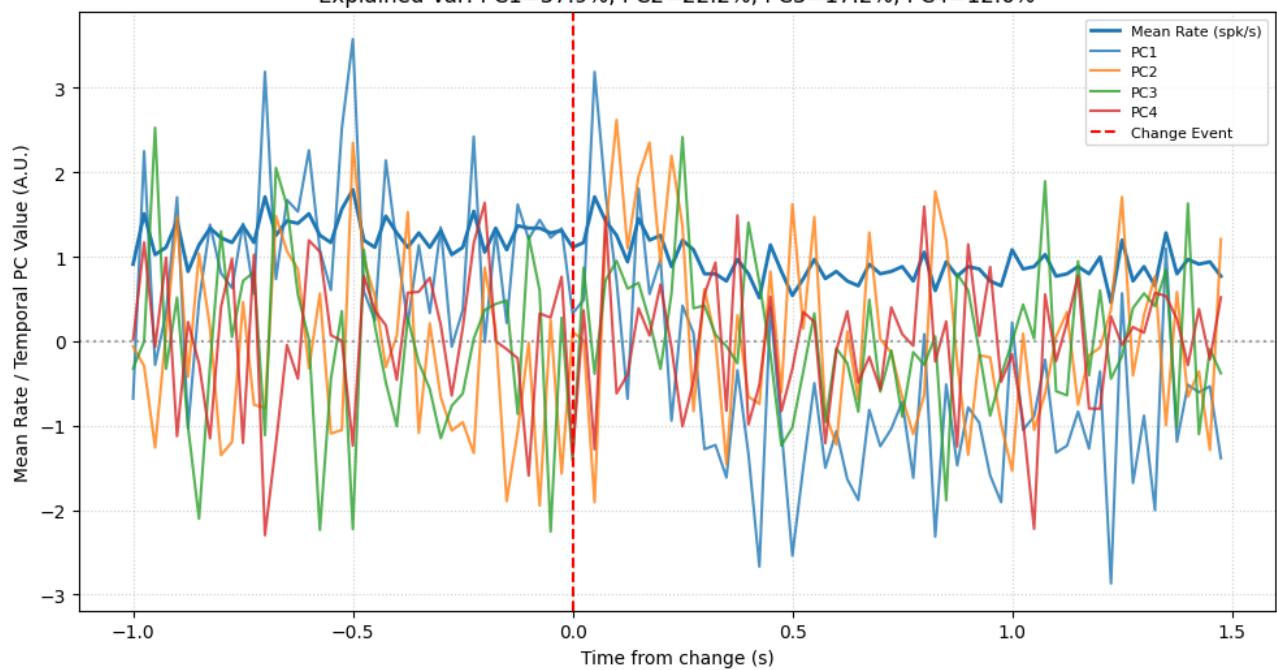




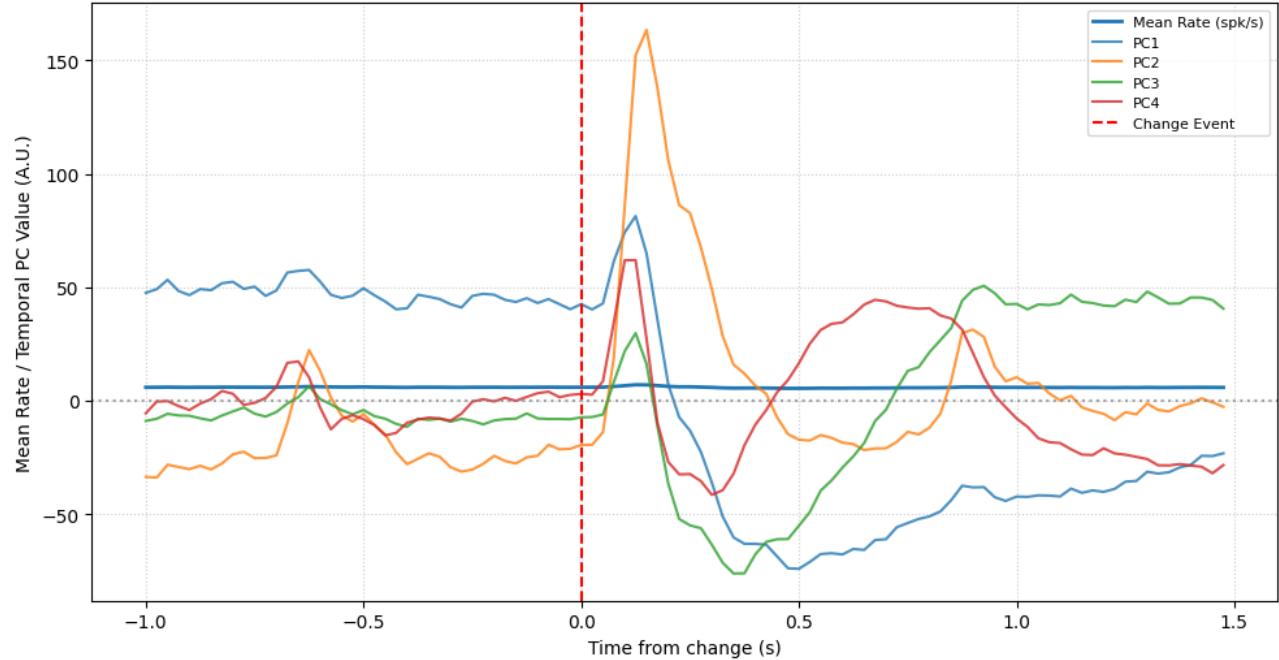
APN | 9565 units | 1,269,480,261 #spikes  
 Explained Var: PC1=32.0%, PC2=11.6%, PC3=7.8%, PC4=4.6%



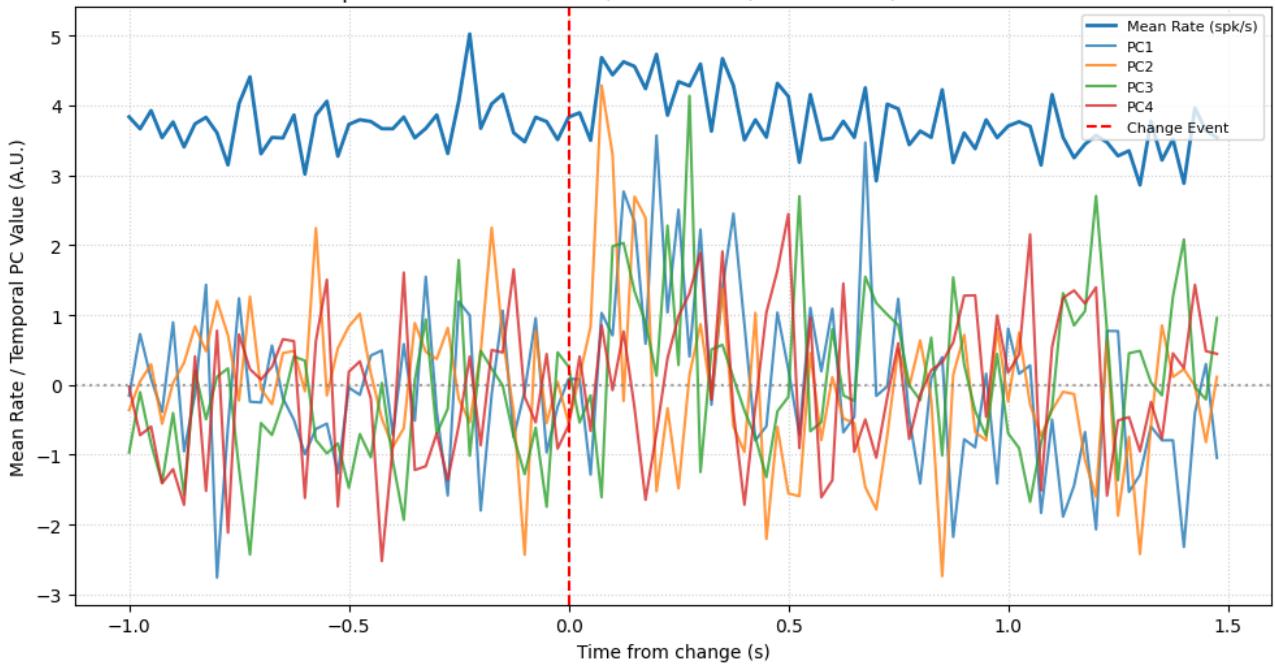
APr | 5 units | 223,780 #spikes  
 Explained Var: PC1=37.9%, PC2=22.2%, PC3=17.2%, PC4=12.6%



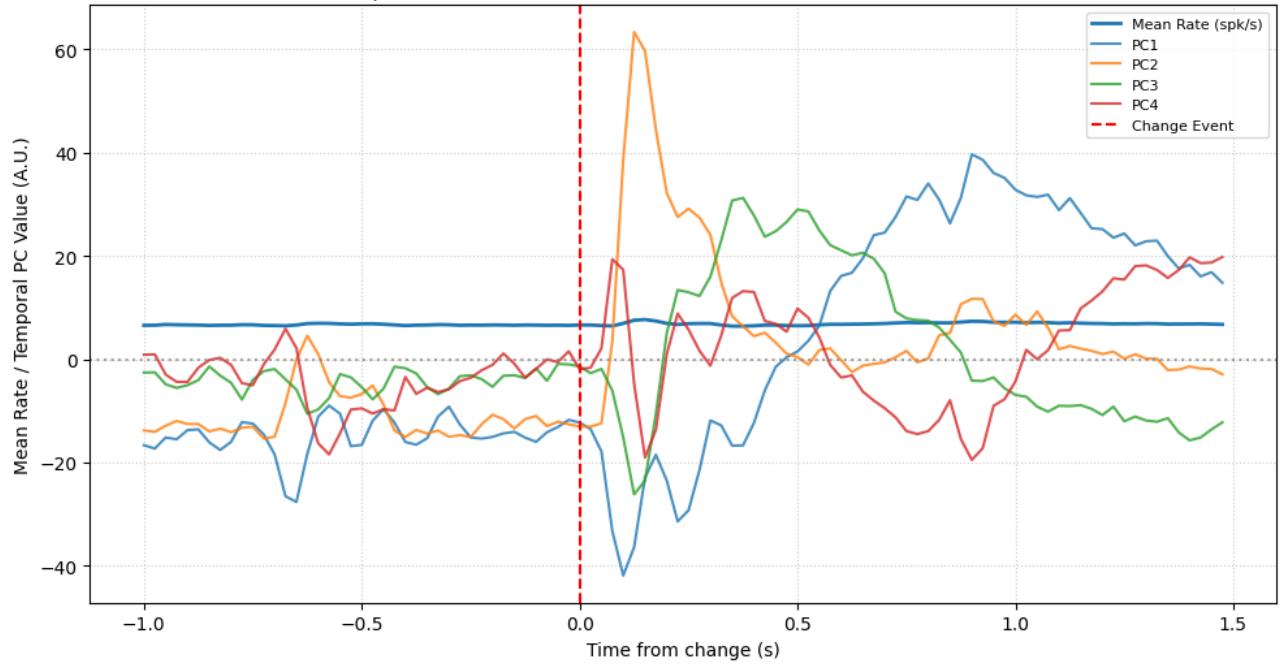
CA1 | 19914 units | 1,209,182,954 #spikes  
 Explained Var: PC1=12.0%, PC2=7.2%, PC3=5.7%, PC4=2.8%



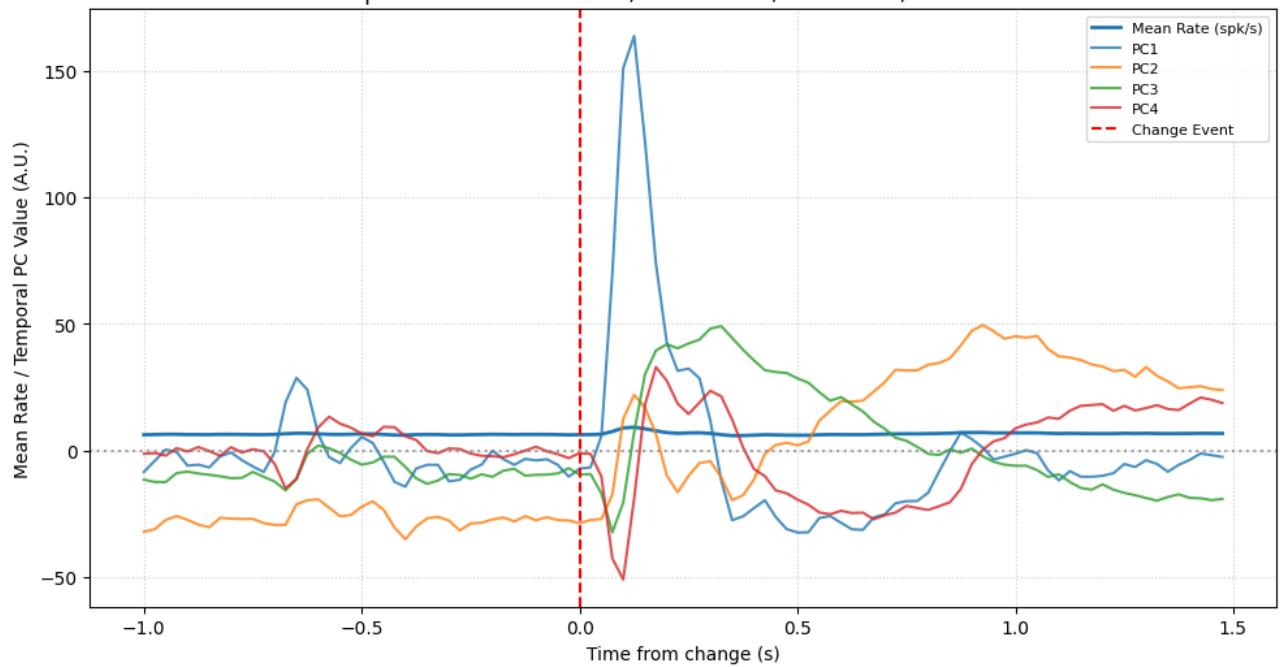
CA2 | 8 units | 411,786 #spikes  
 Explained Var: PC1=17.8%, PC2=15.9%, PC3=15.4%, PC4=14.0%

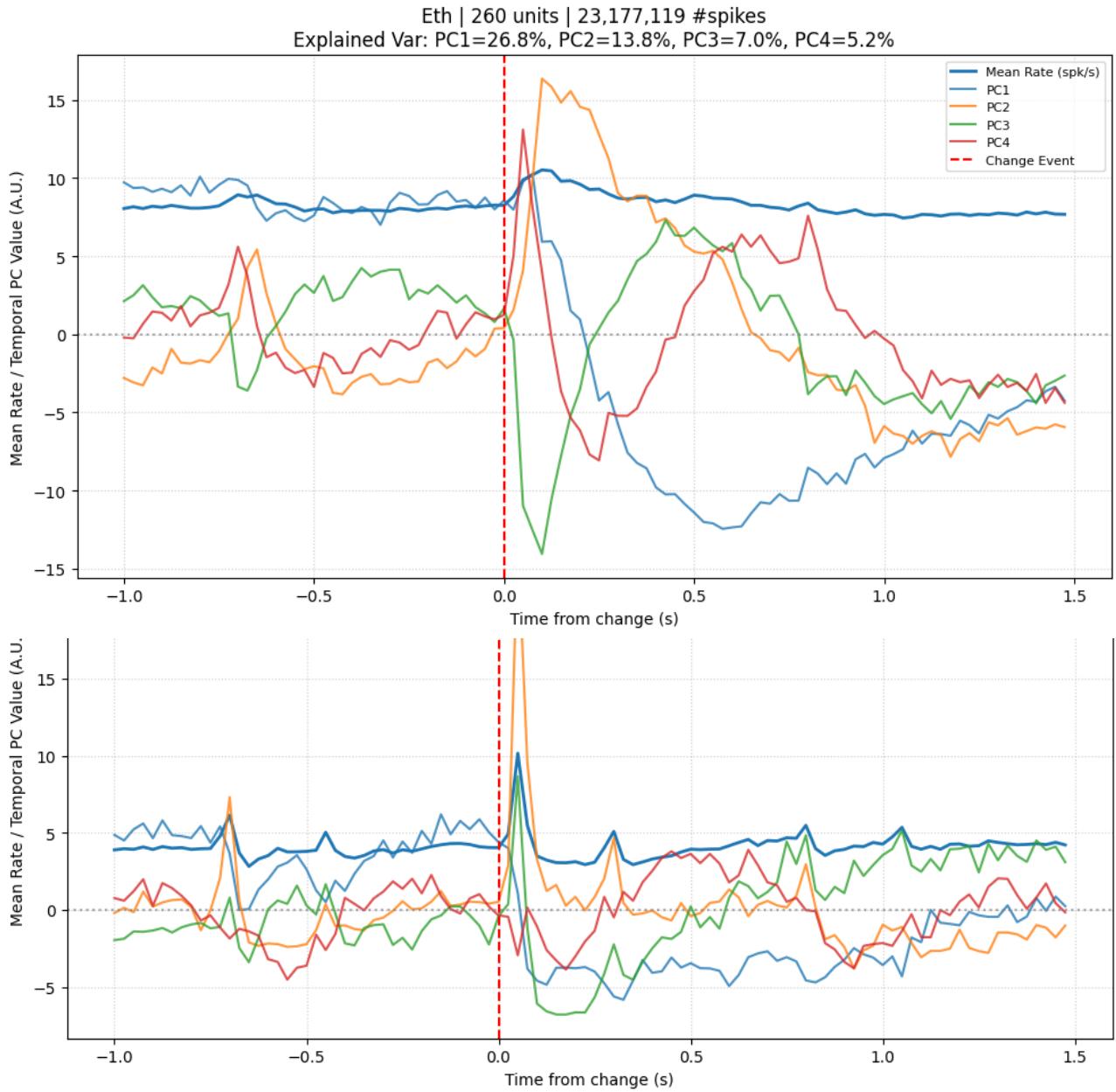


CA3 | 3474 units | 235,841,453 #spikes  
Explained Var: PC1=13.3%, PC2=6.5%, PC3=4.6%, PC4=2.9%

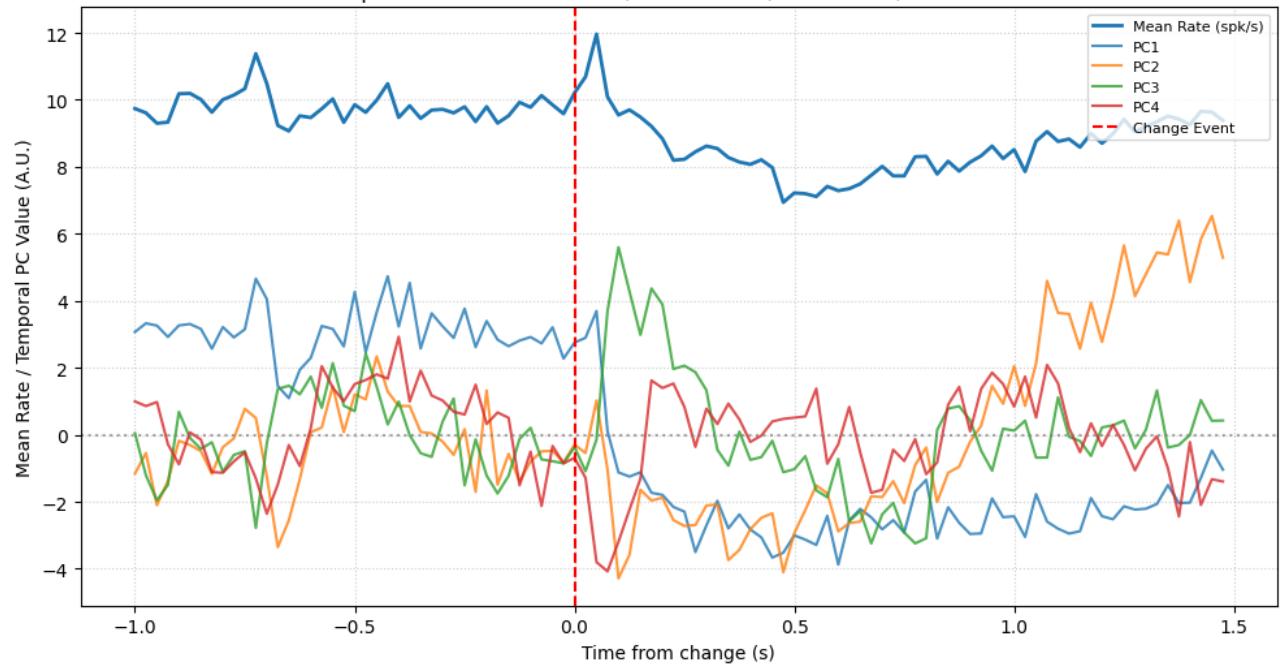


DG | 6842 units | 448,034,239 #spikes  
Explained Var: PC1=14.3%, PC2=11.3%, PC3=5.3%, PC4=3.6%

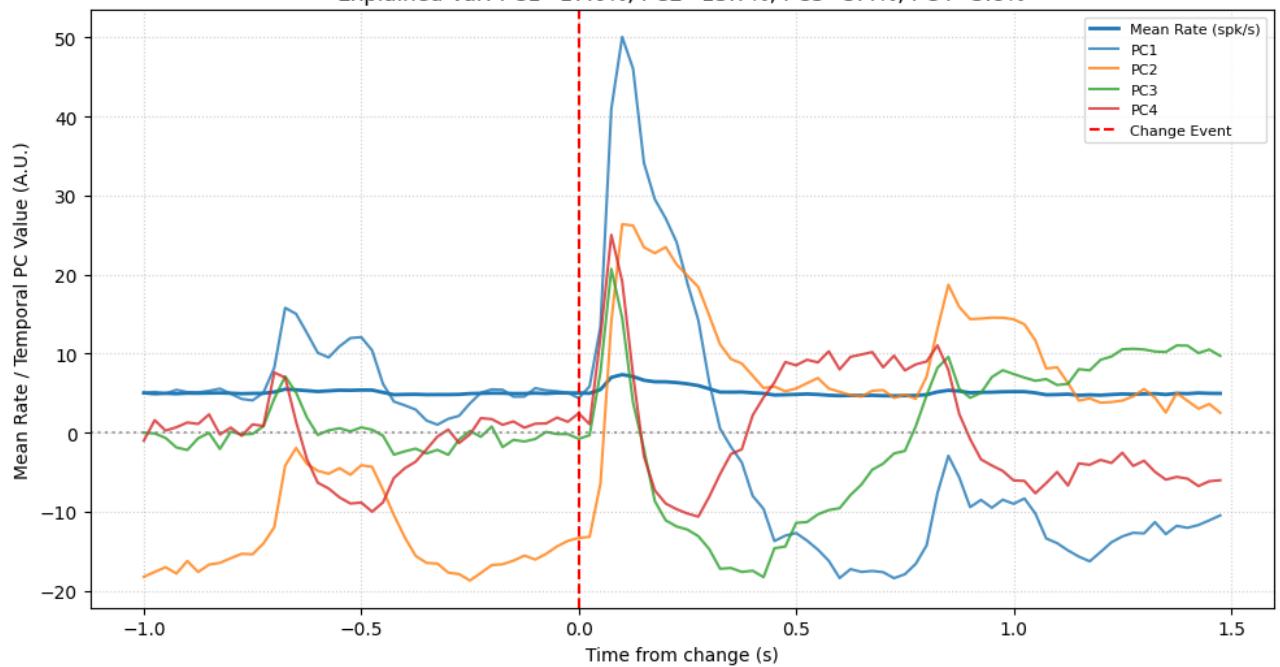


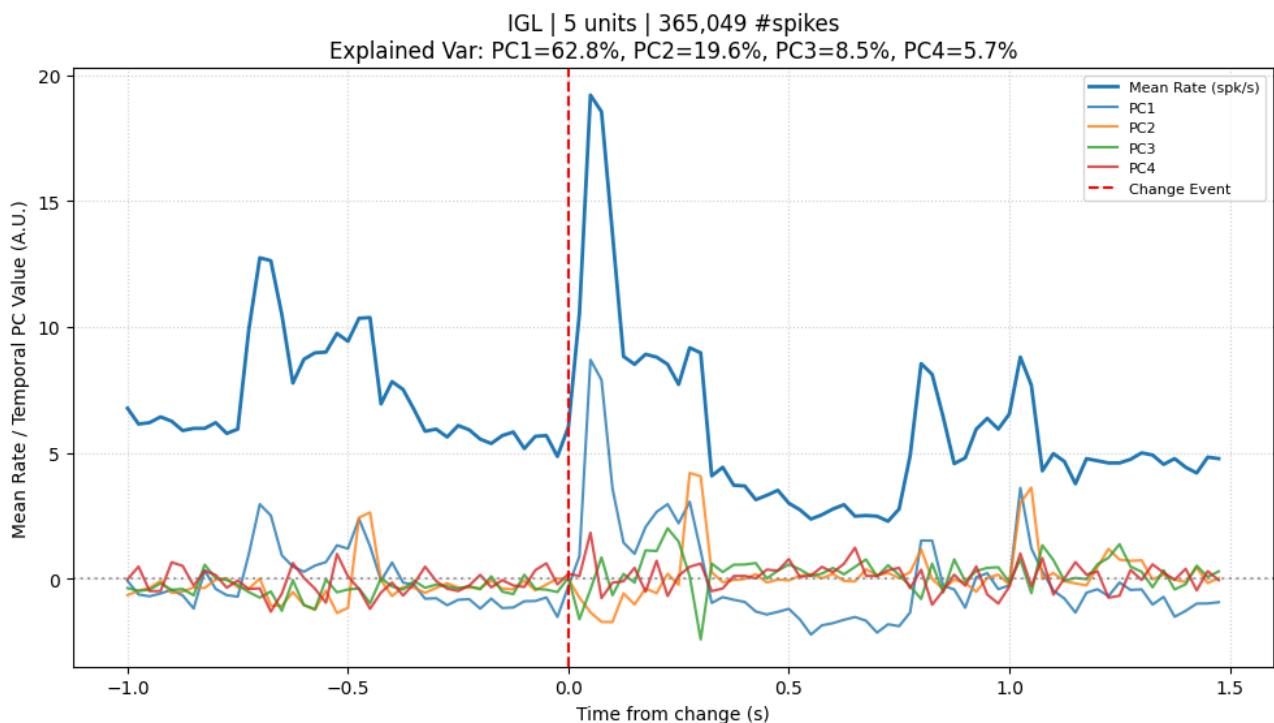
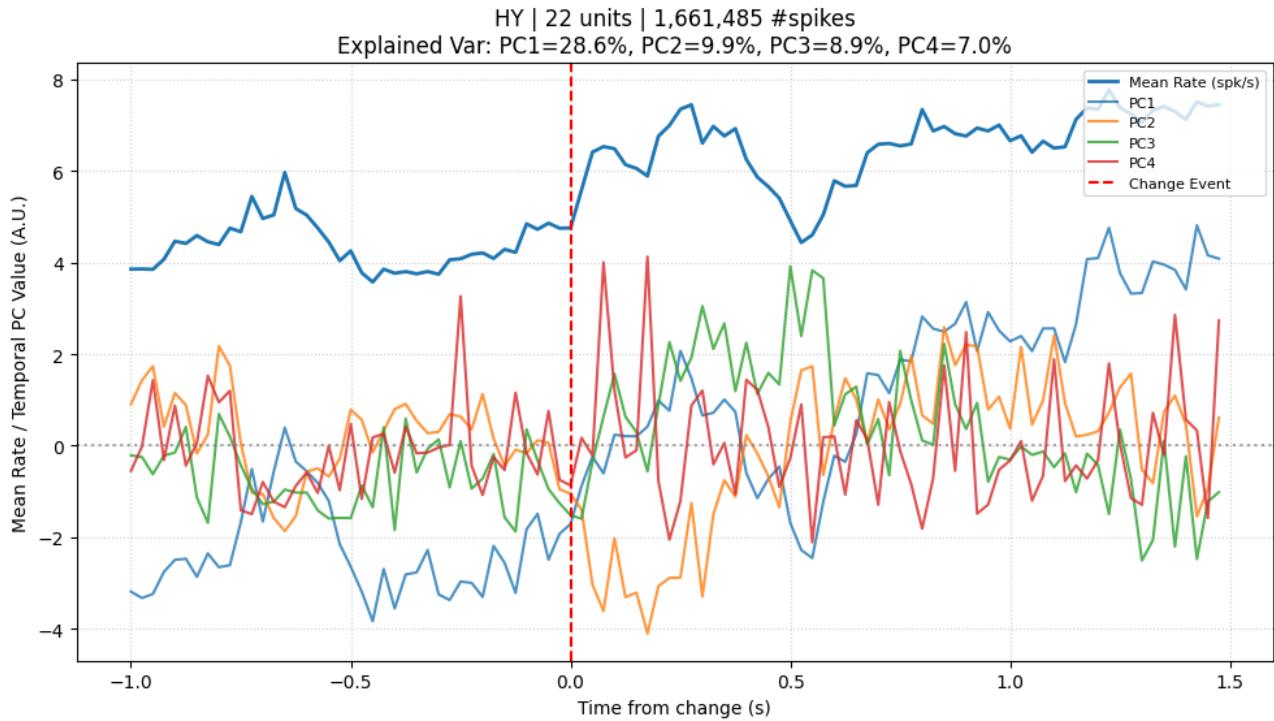


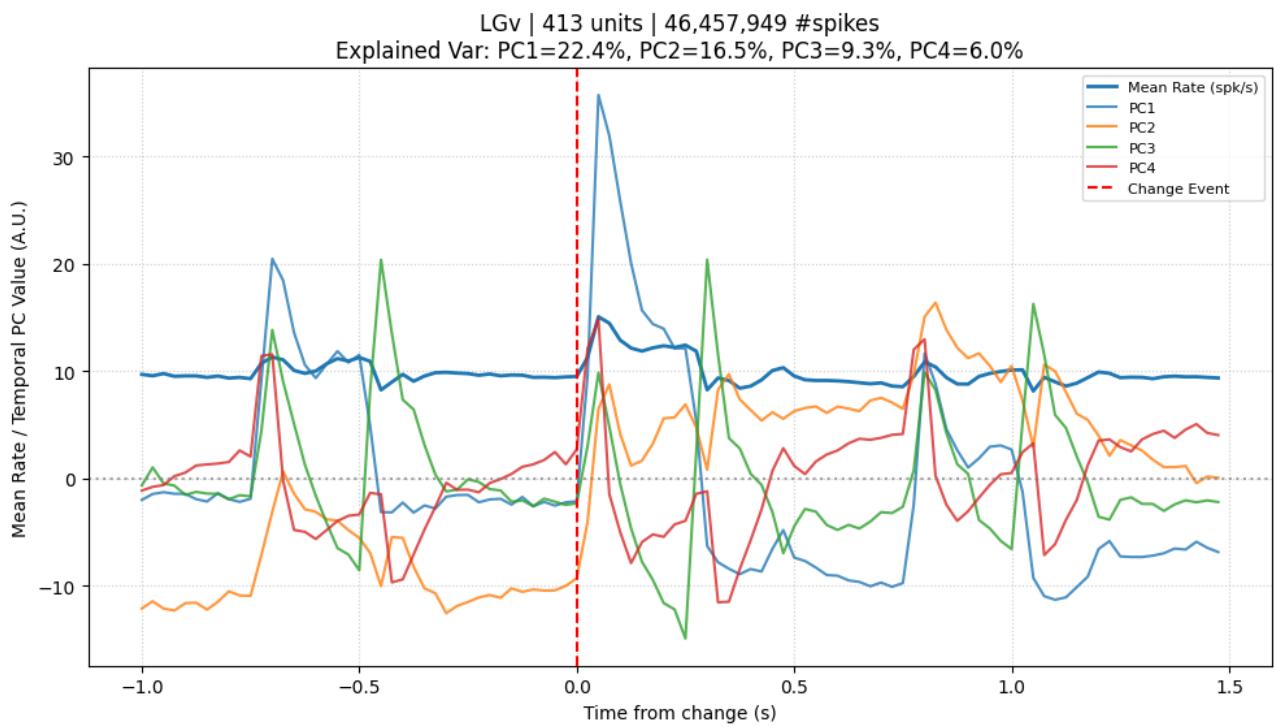
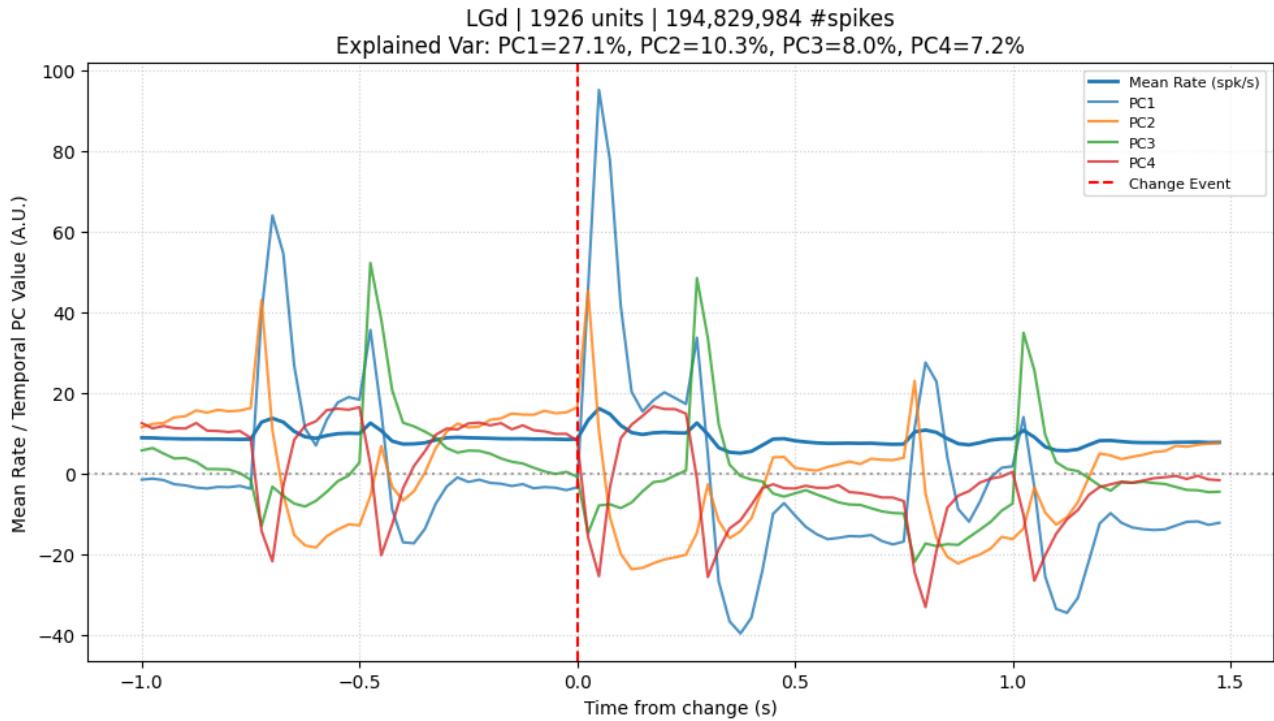
FF | 31 units | 2,458,316 #spikes  
Explained Var: PC1=24.9%, PC2=21.2%, PC3=8.3%, PC4=5.6%

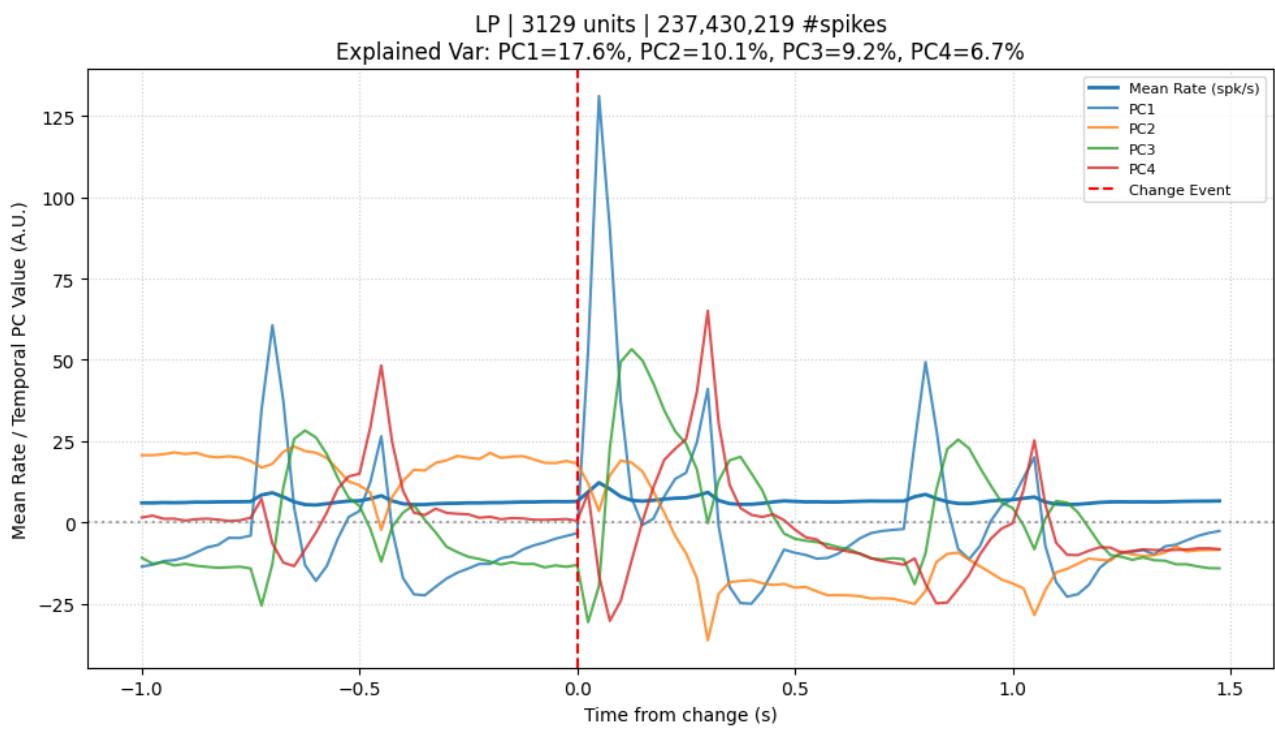


HPF | 1204 units | 82,053,056 #spikes  
Explained Var: PC1=17.0%, PC2=13.7%, PC3=5.4%, PC4=3.8%

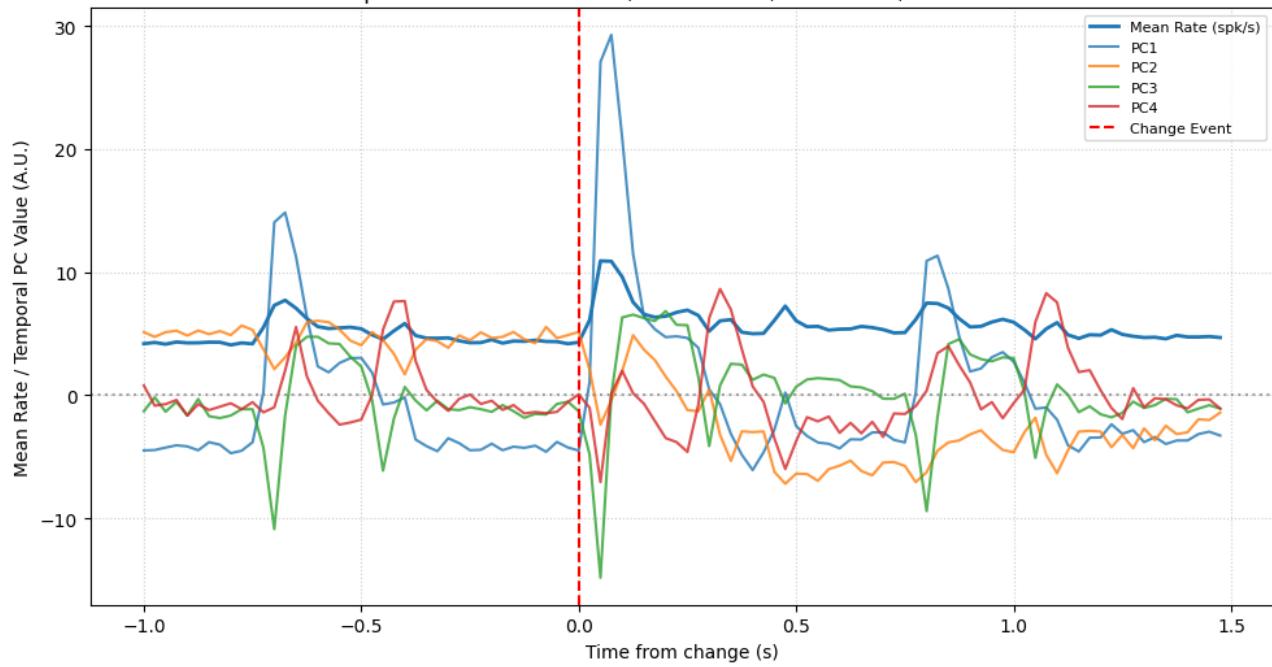




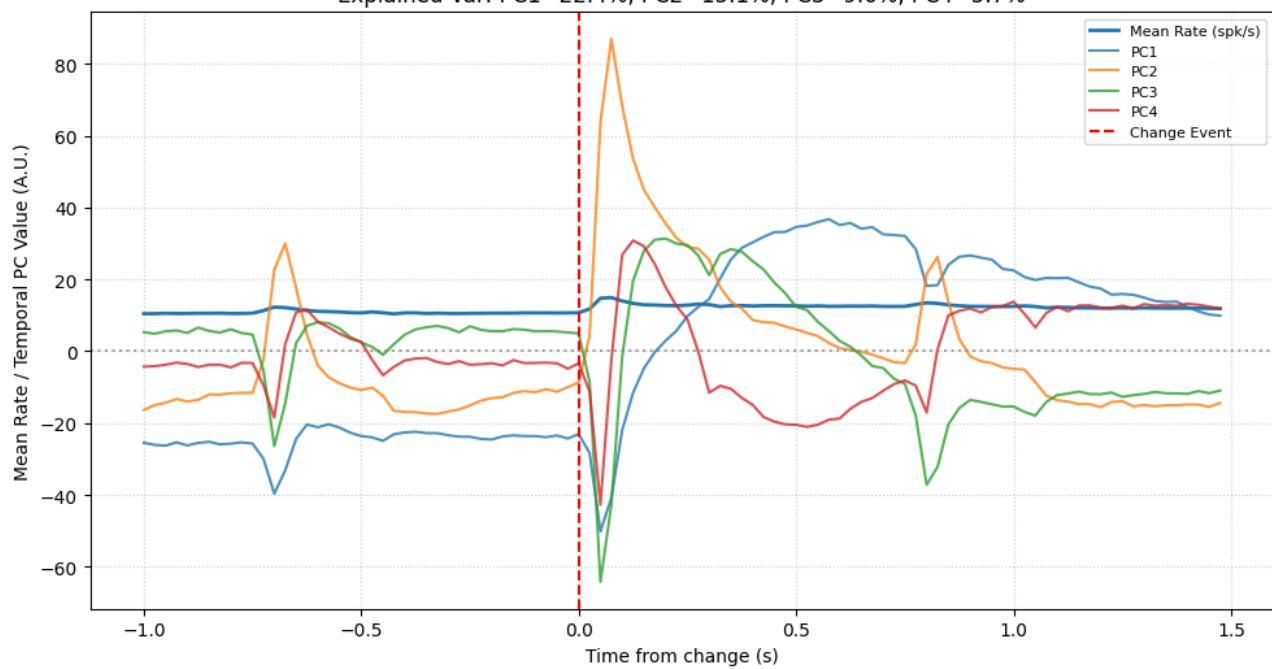


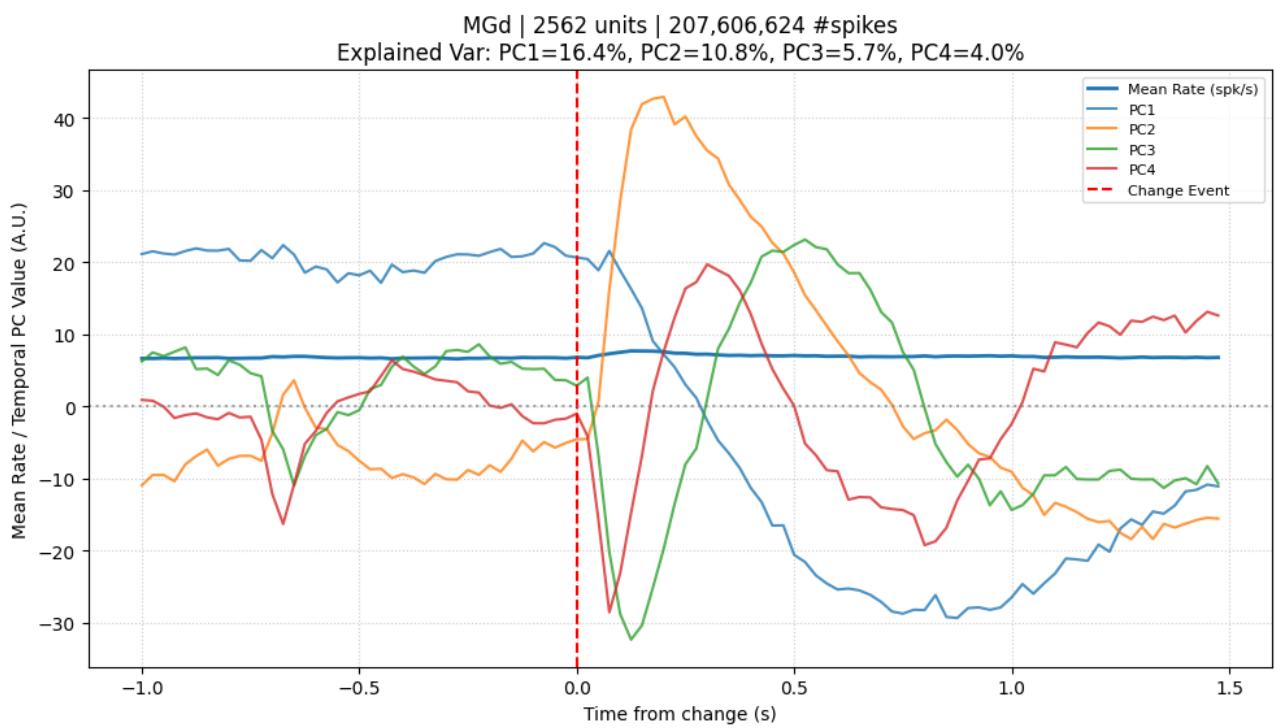


LT | 164 units | 11,408,785 #spikes  
 Explained Var: PC1=26.9%, PC2=12.6%, PC3=7.2%, PC4=5.5%

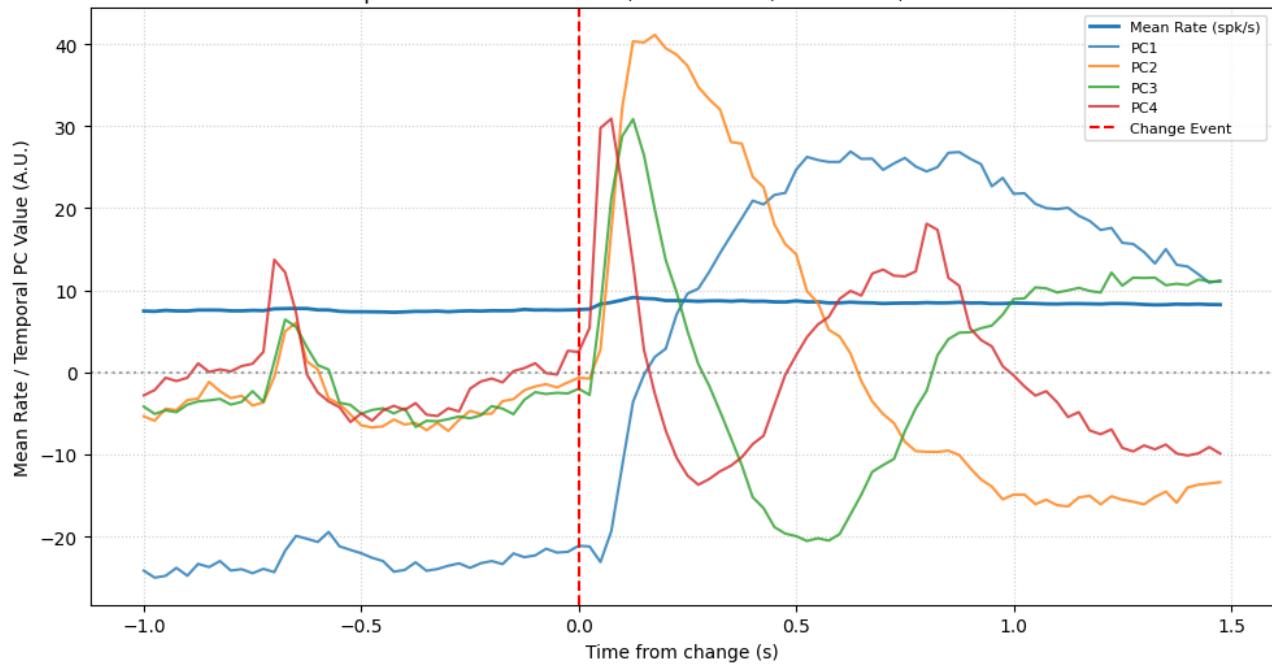


MB | 2818 units | 321,518,979 #spikes  
 Explained Var: PC1=22.4%, PC2=15.1%, PC3=9.6%, PC4=5.7%

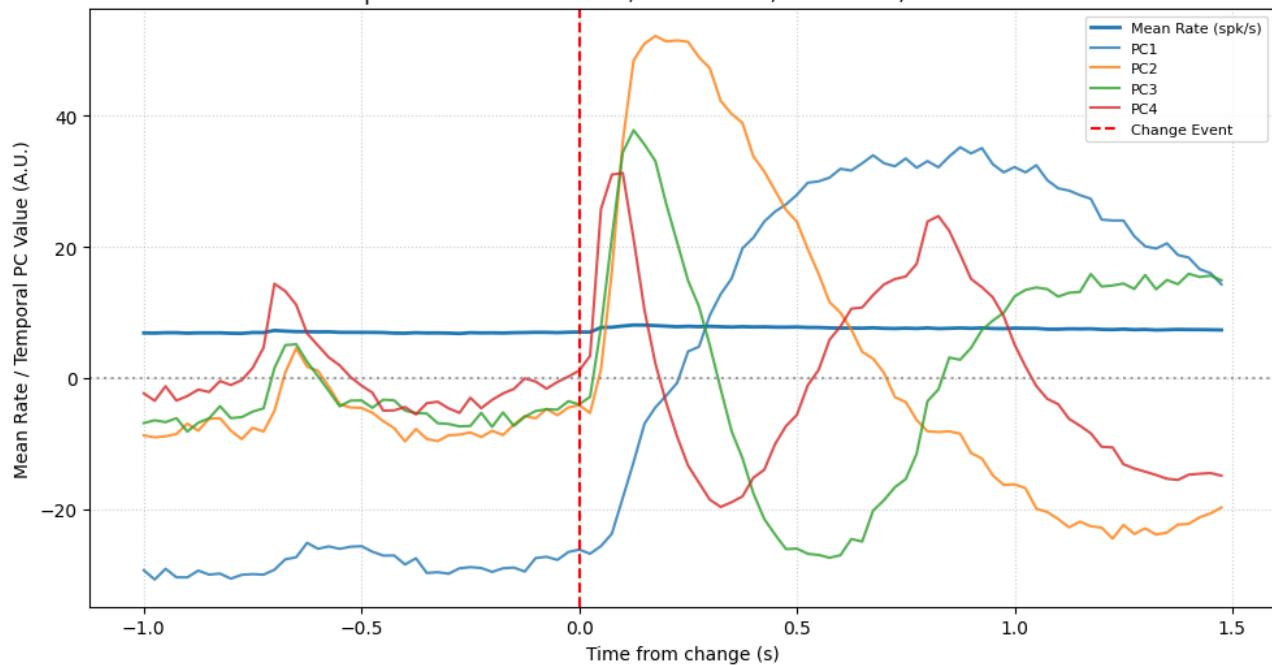


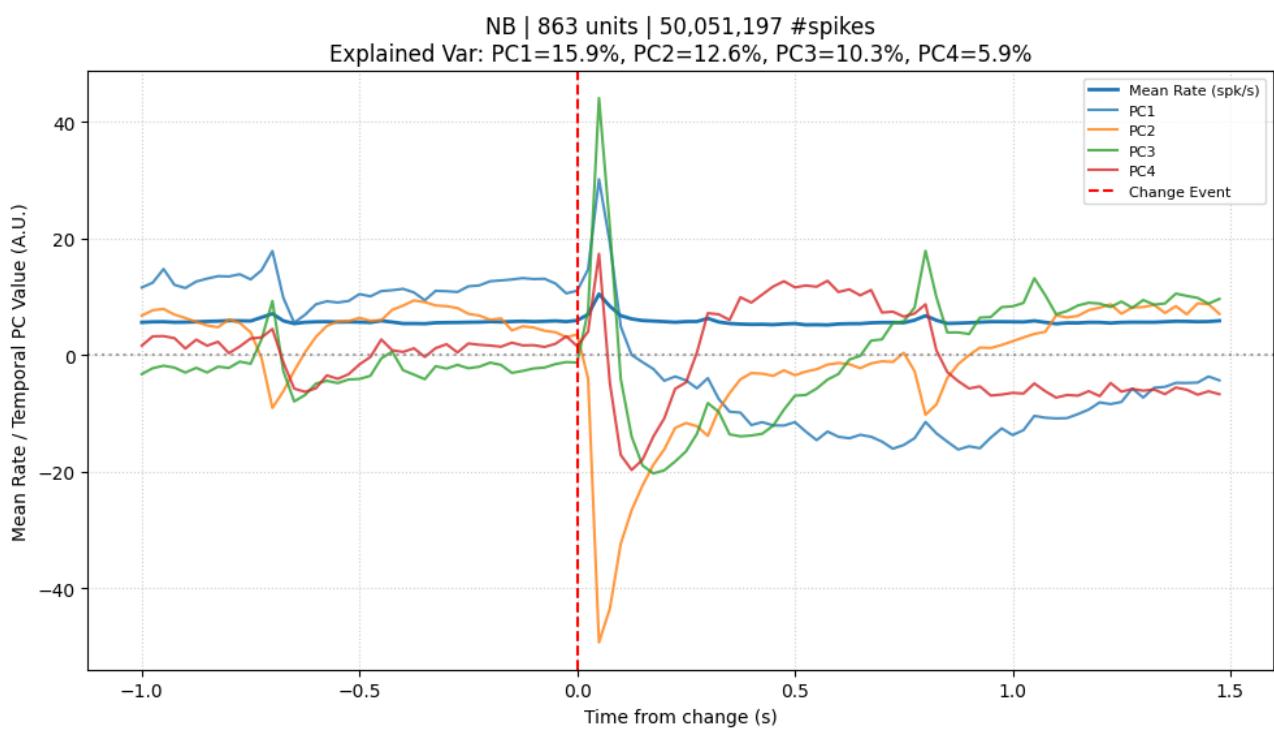
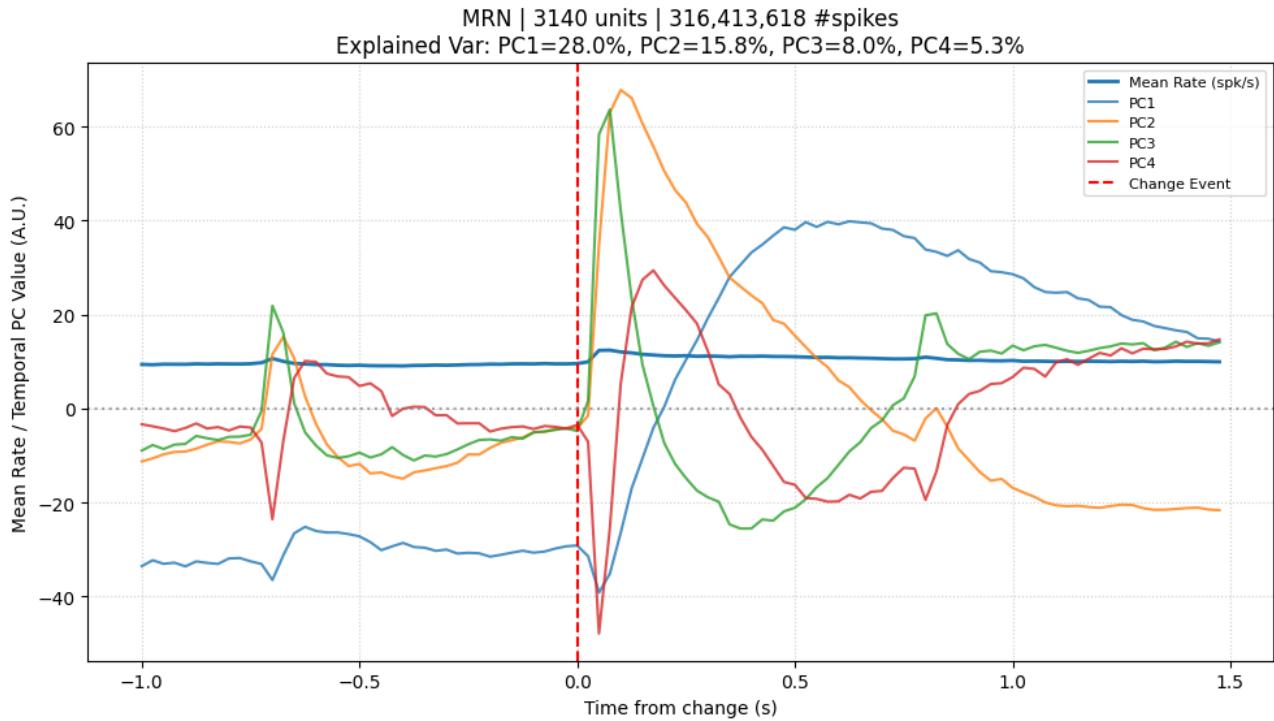


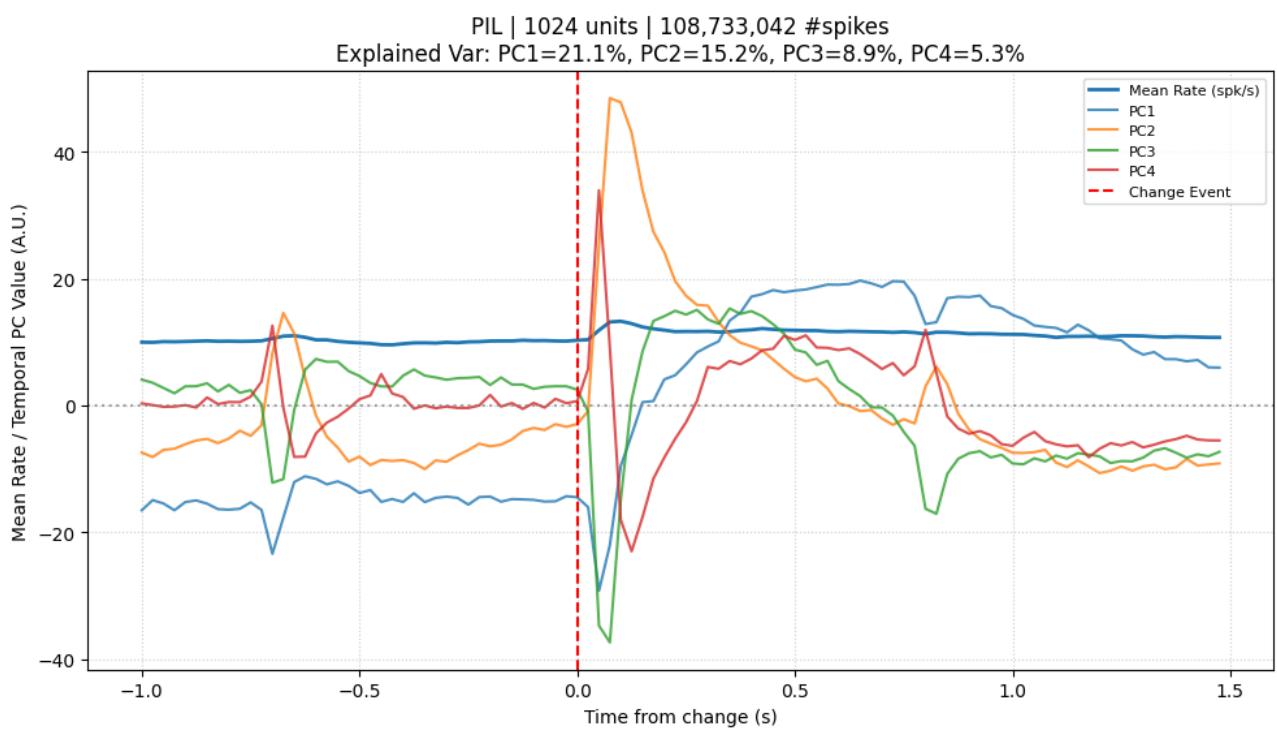
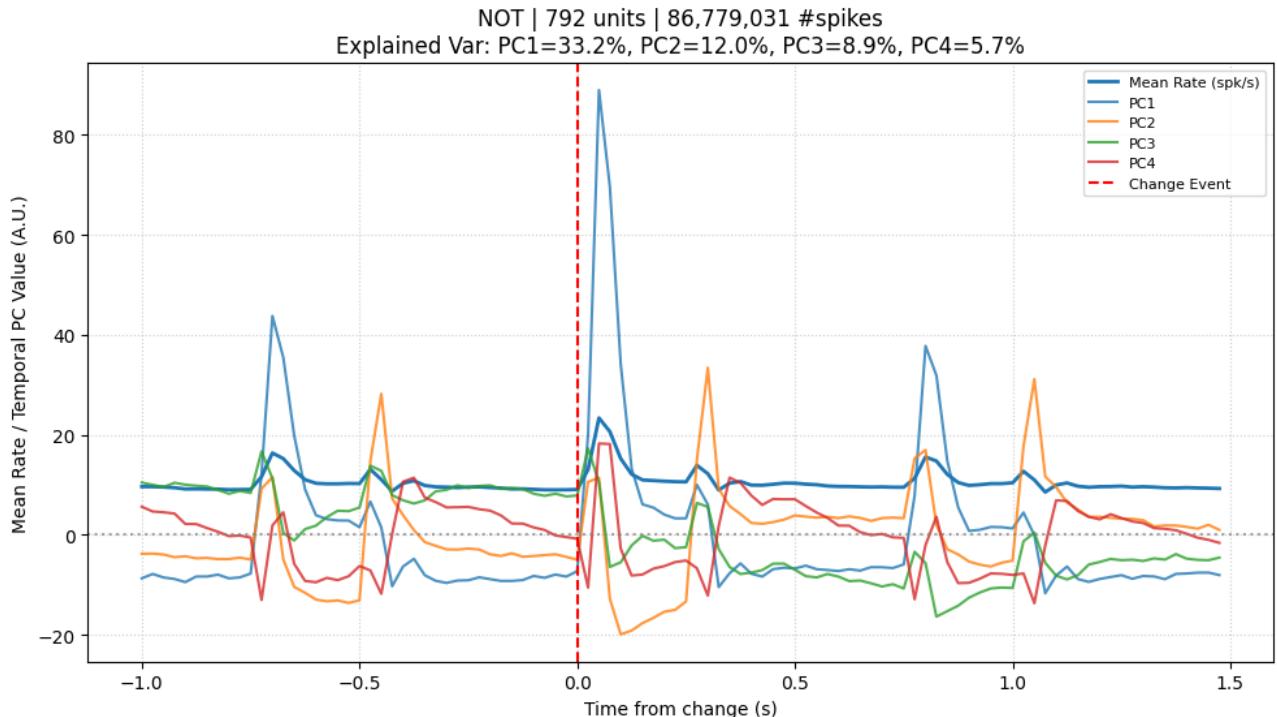
MGm | 1952 units | 161,160,391 #spikes  
 Explained Var: PC1=23.1%, PC2=12.6%, PC3=5.8%, PC4=3.9%

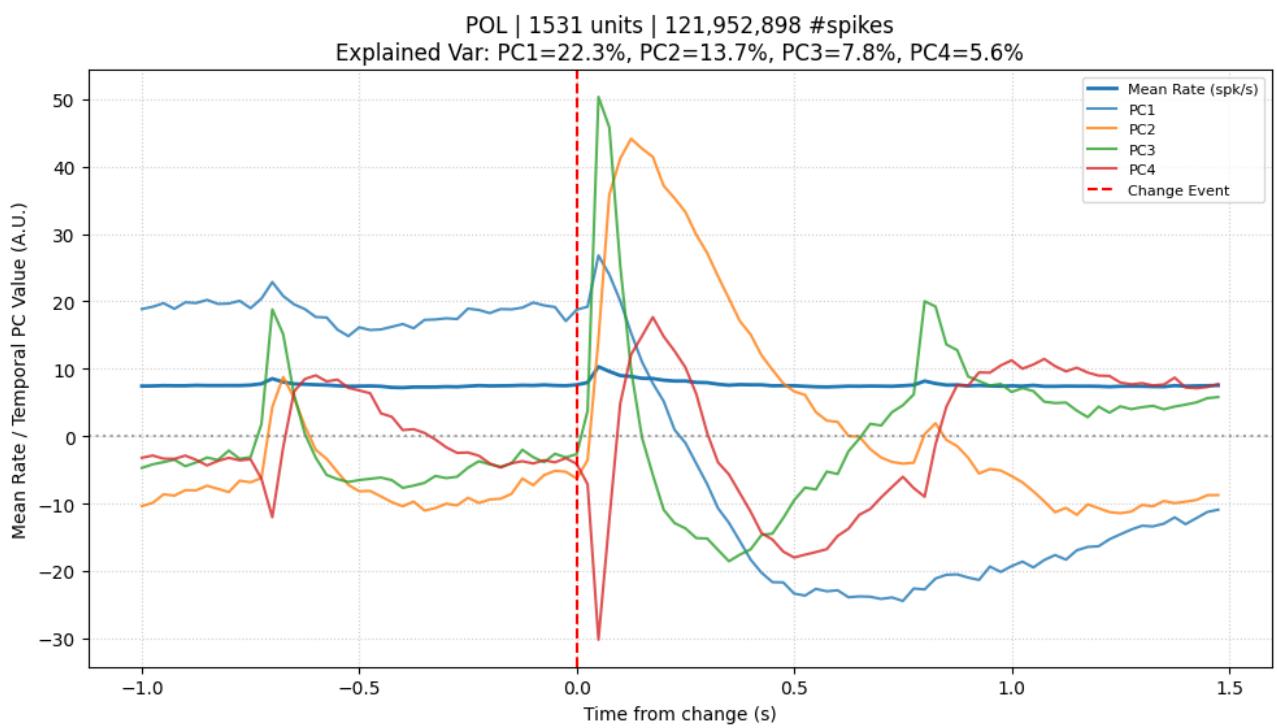


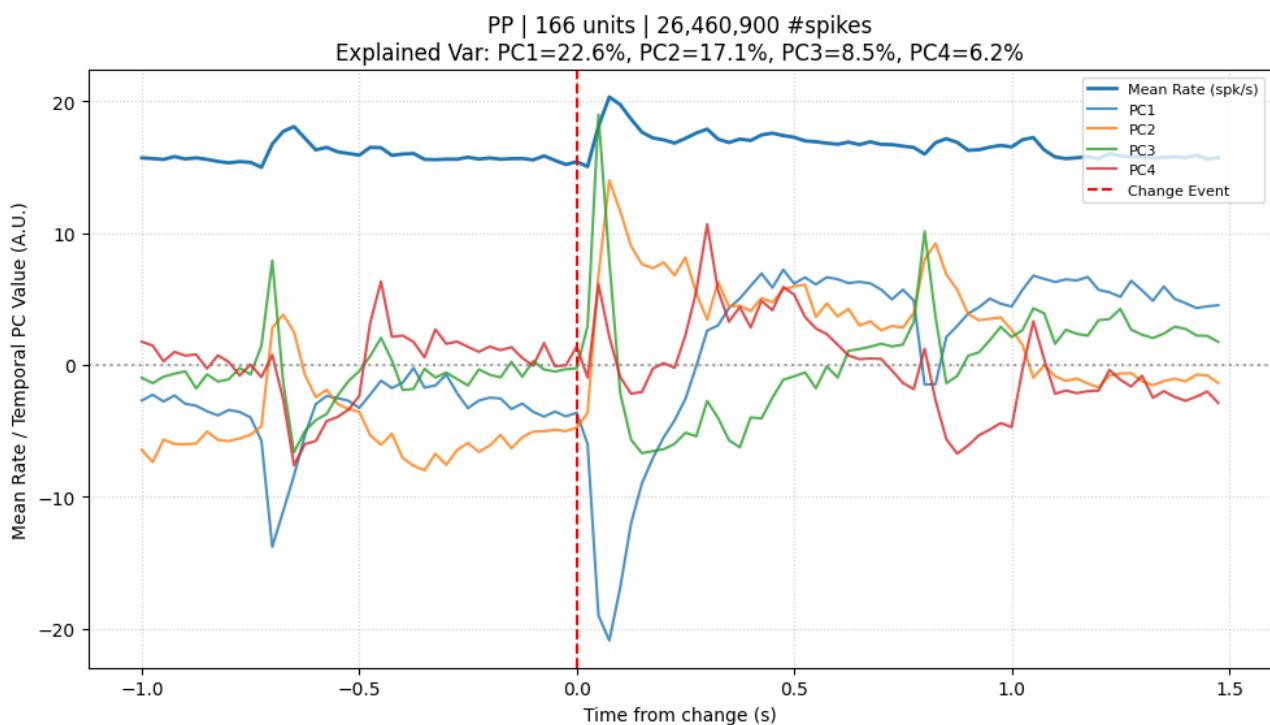
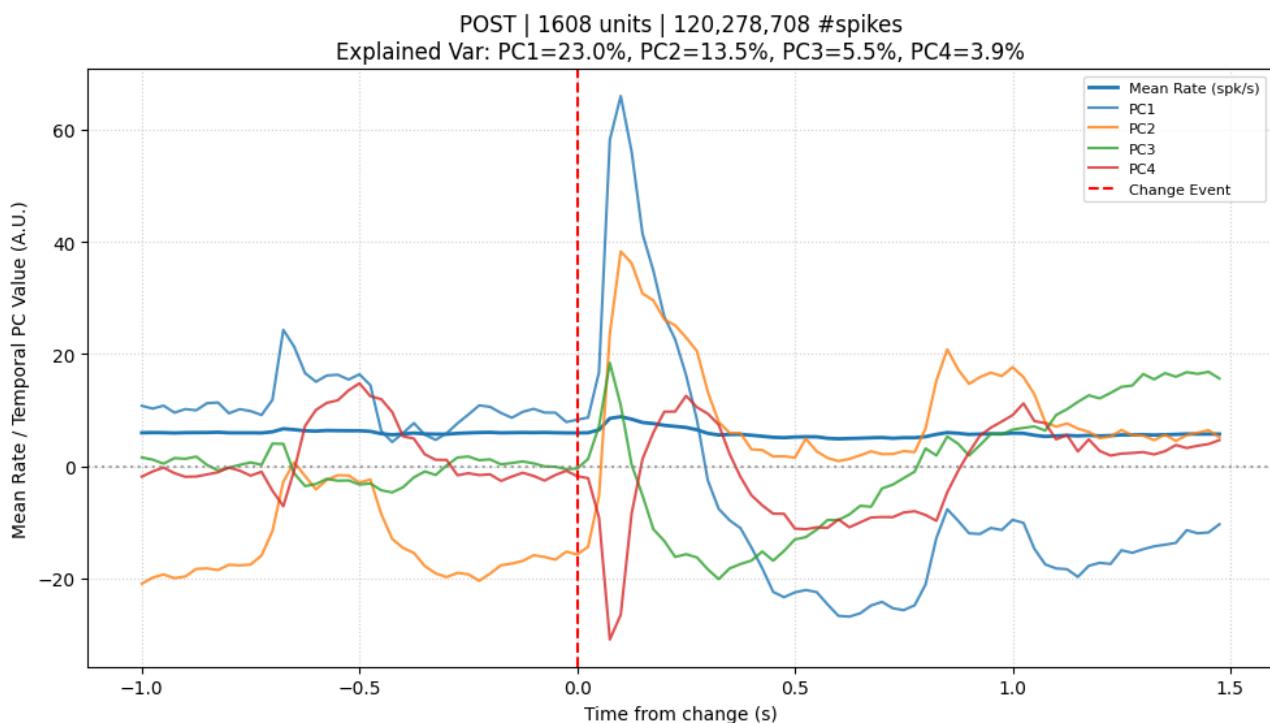
MGv | 3857 units | 316,592,232 #spikes  
 Explained Var: PC1=18.9%, PC2=11.7%, PC3=5.5%, PC4=3.5%

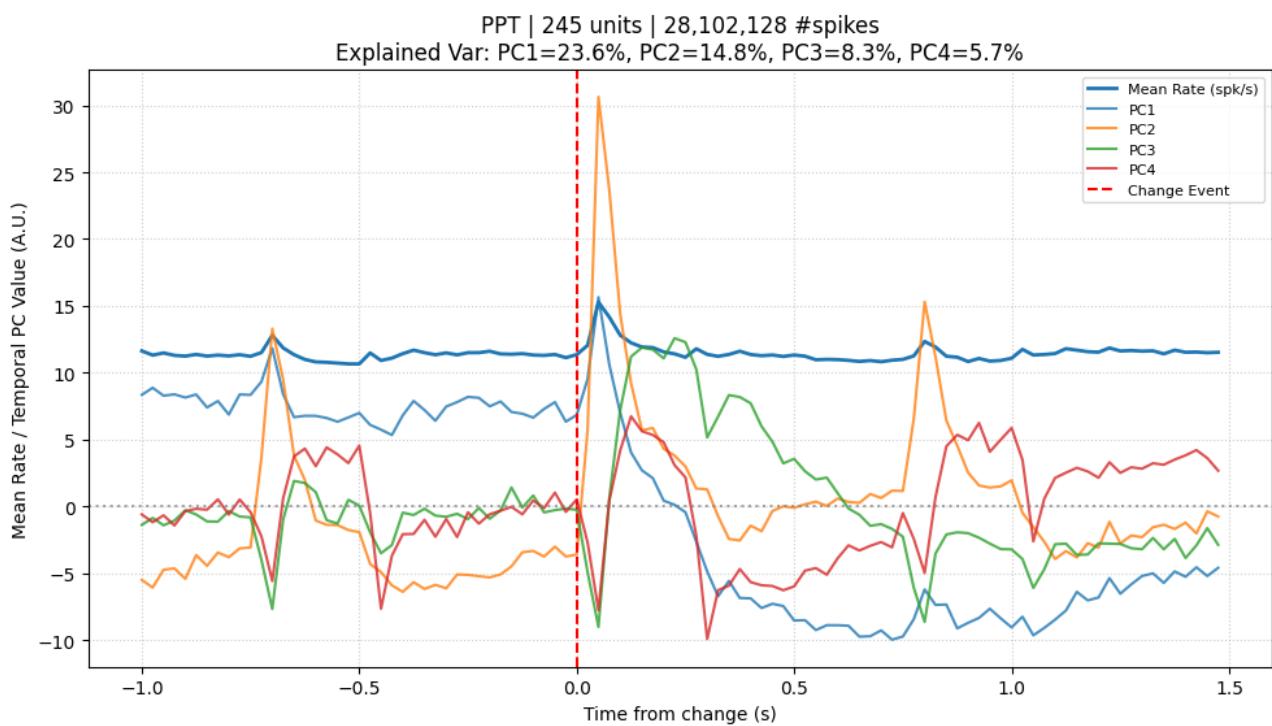




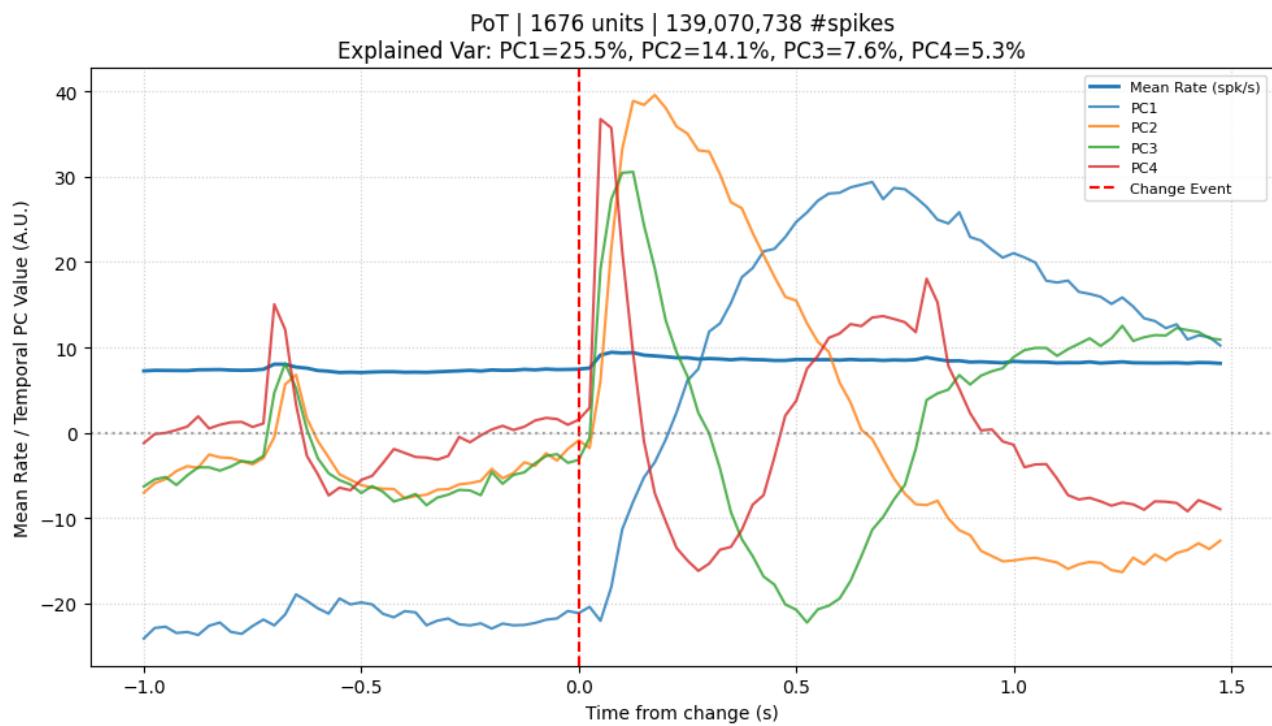
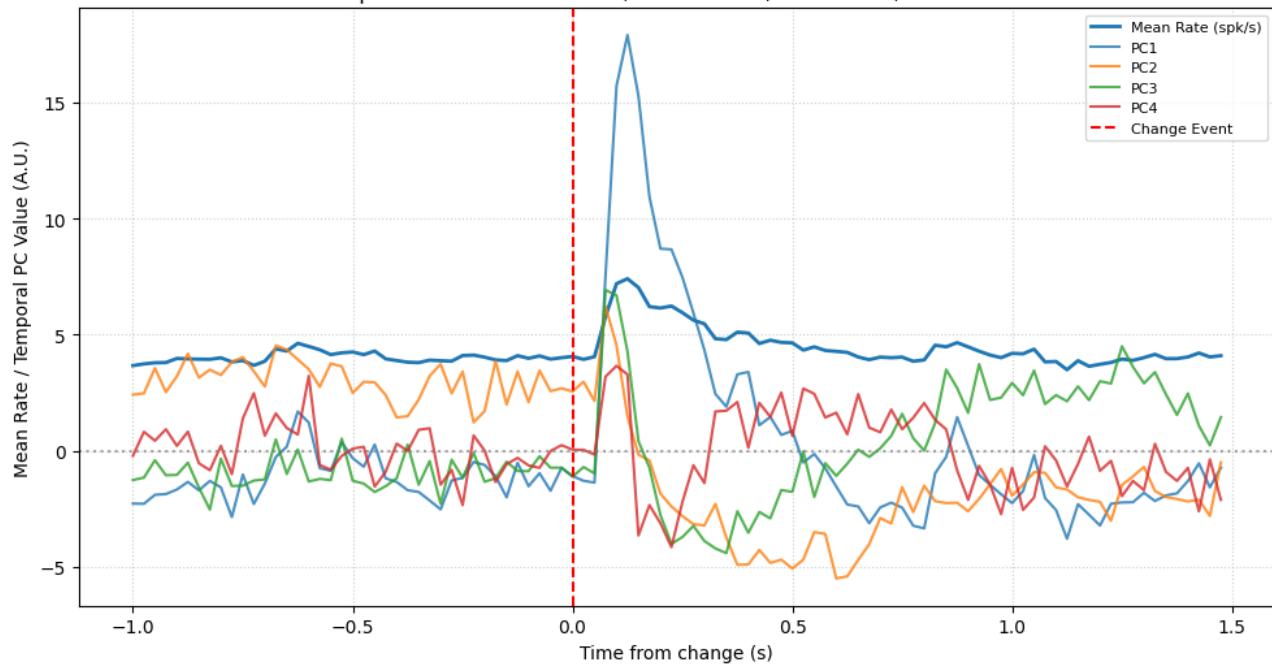




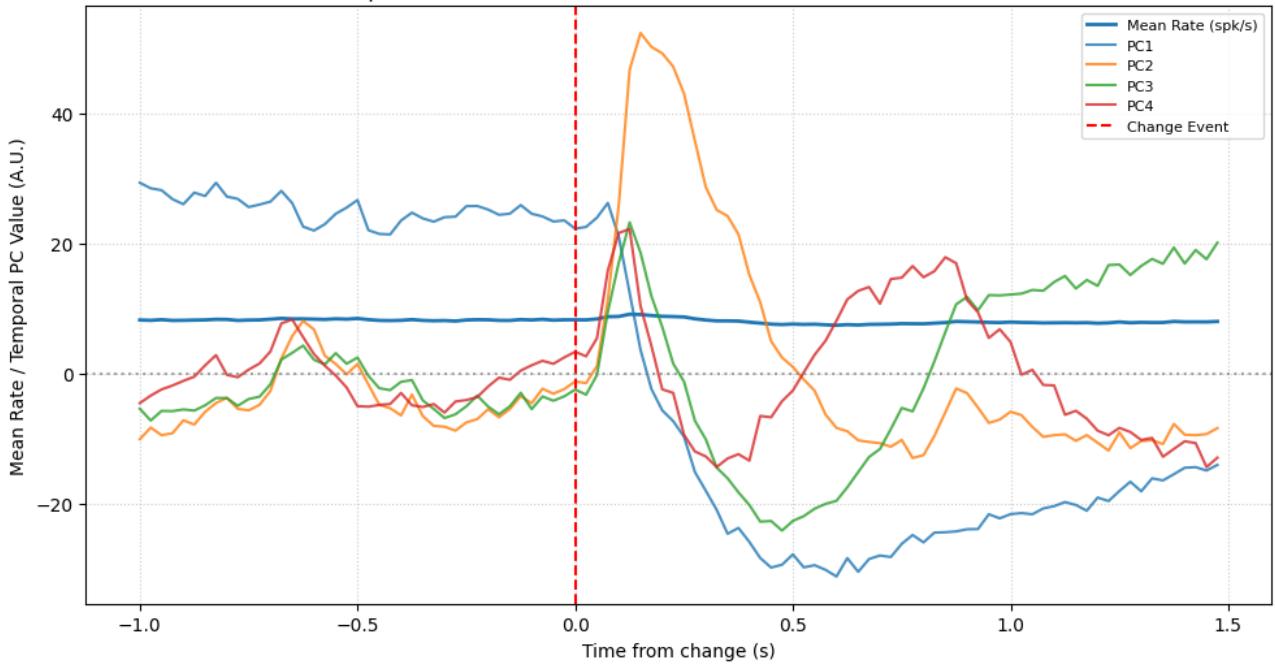




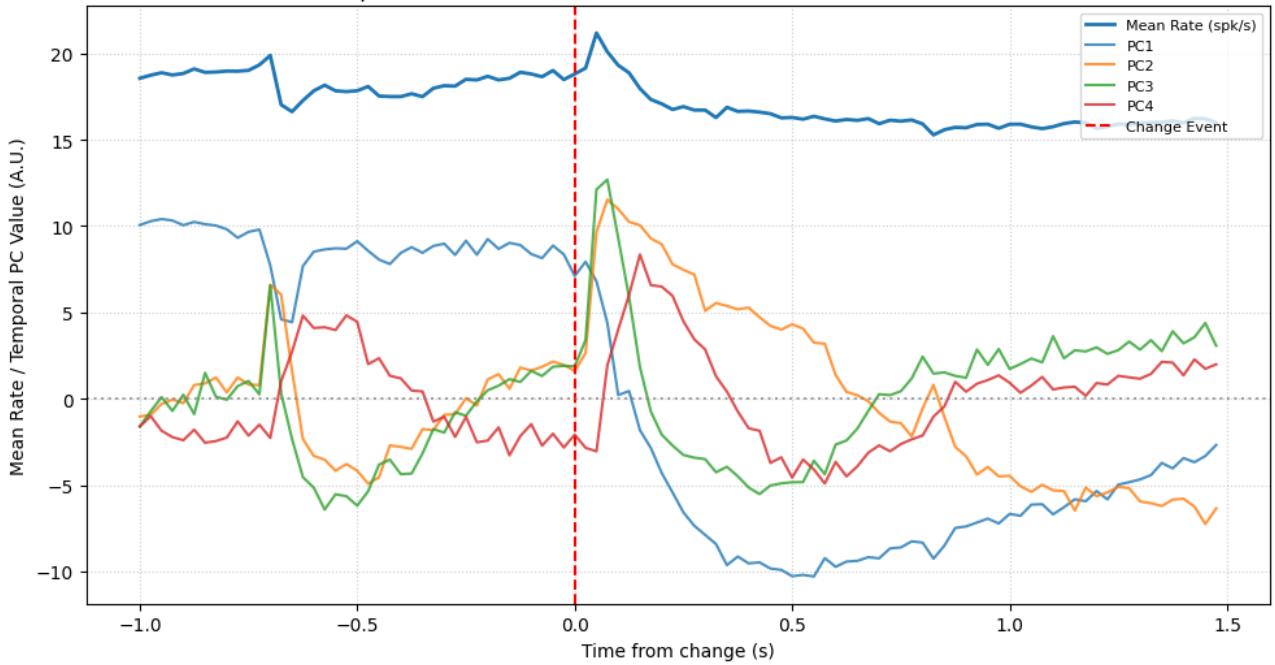
PRE | 71 units | 4,170,580 #spikes  
 Explained Var: PC1=21.9%, PC2=13.4%, PC3=7.3%, PC4=3.6%



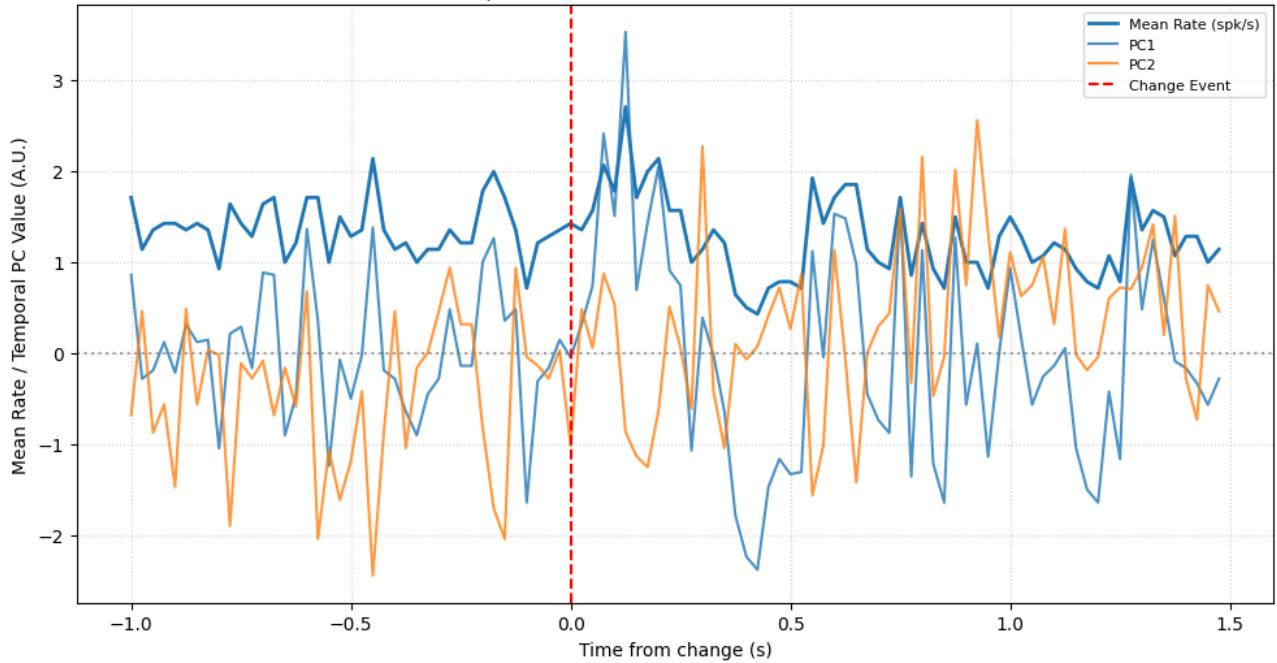
ProS | 2745 units | 238,939,645 #spikes  
 Explained Var: PC1=20.4%, PC2=8.6%, PC3=5.2%, PC4=2.7%



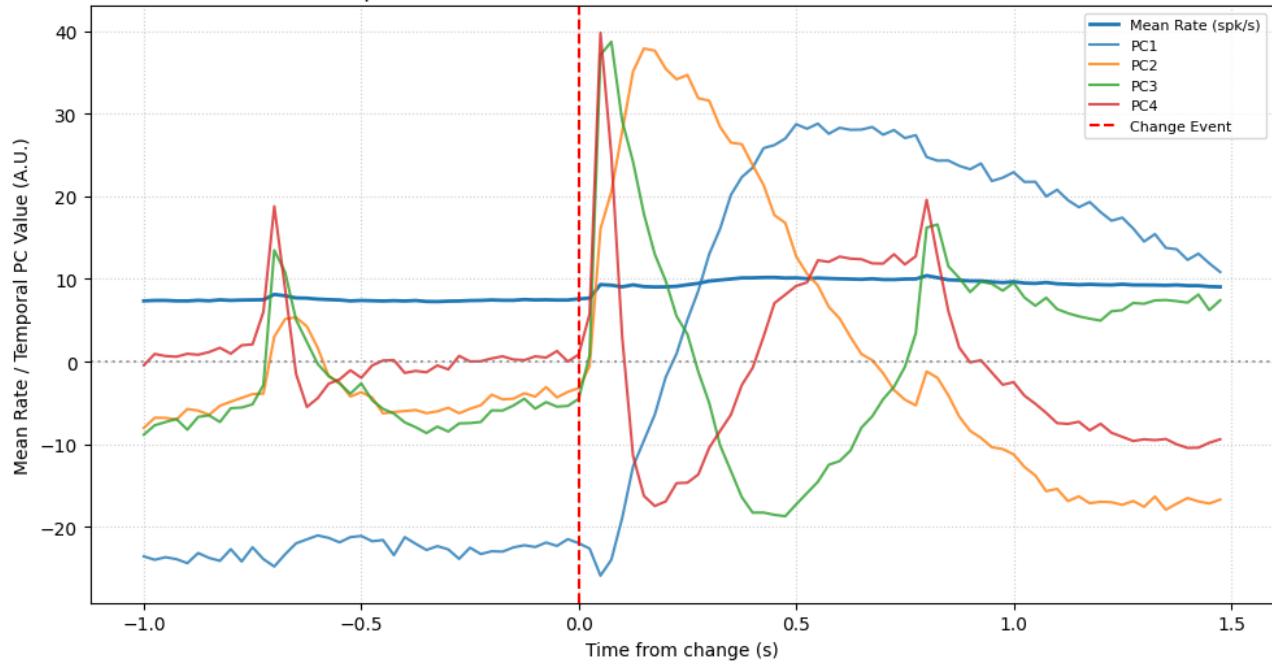
RN | 161 units | 30,851,455 #spikes  
 Explained Var: PC1=38.9%, PC2=13.4%, PC3=8.3%, PC4=5.0%



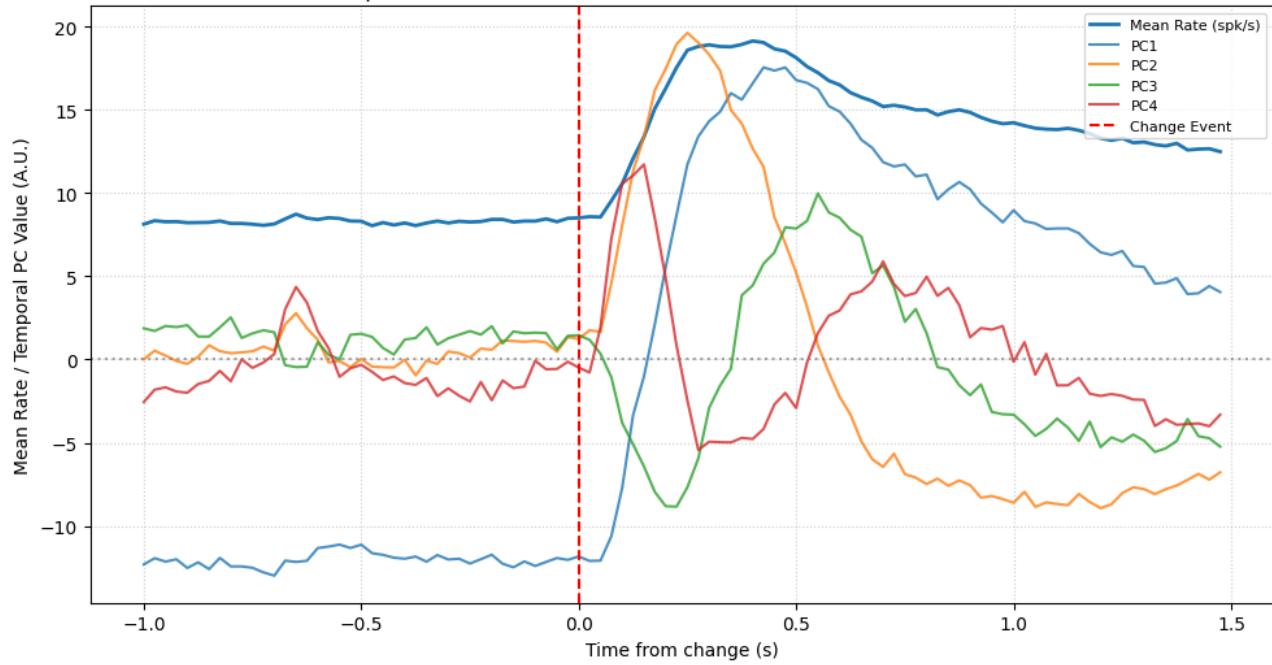
RSPd | 2 units | 107,216 #spikes  
Explained Var: PC1=53.6%, PC2=46.4%

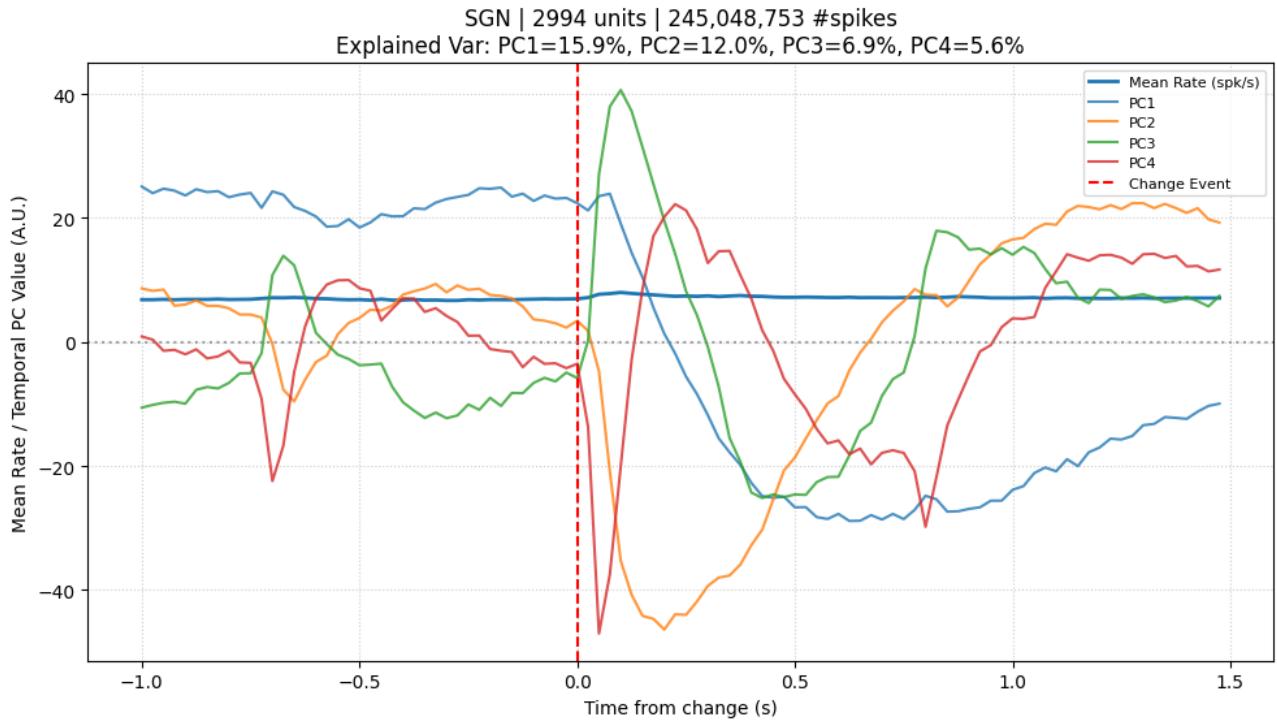


SCig | 1613 units | 125,873,608 #spikes  
 Explained Var: PC1=29.6%, PC2=13.9%, PC3=7.5%, PC4=5.3%

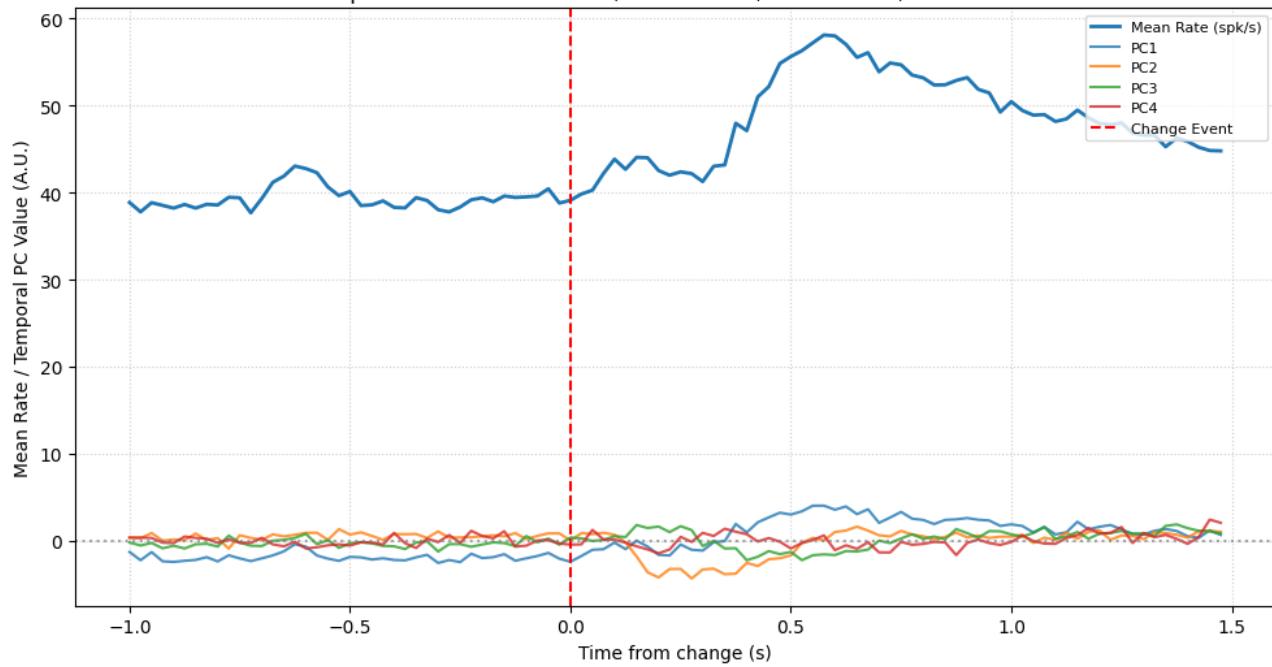


SCiw | 287 units | 24,354,544 #spikes  
 Explained Var: PC1=45.4%, PC2=19.9%, PC3=6.0%, PC4=4.3%

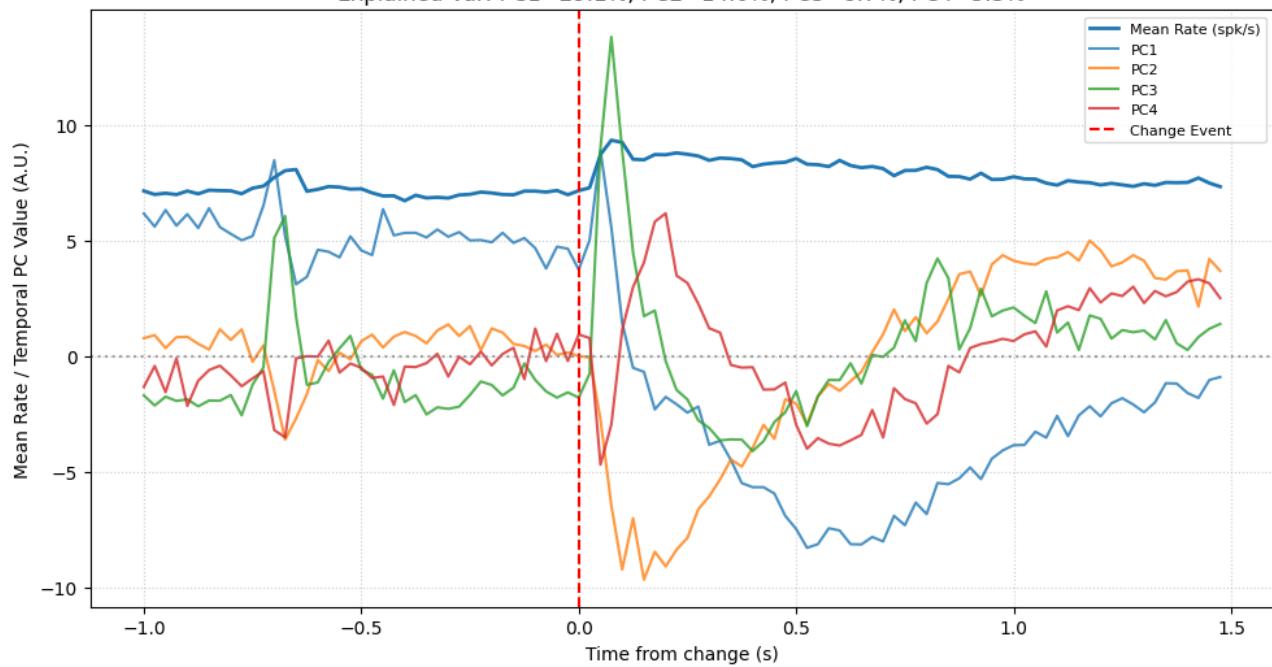




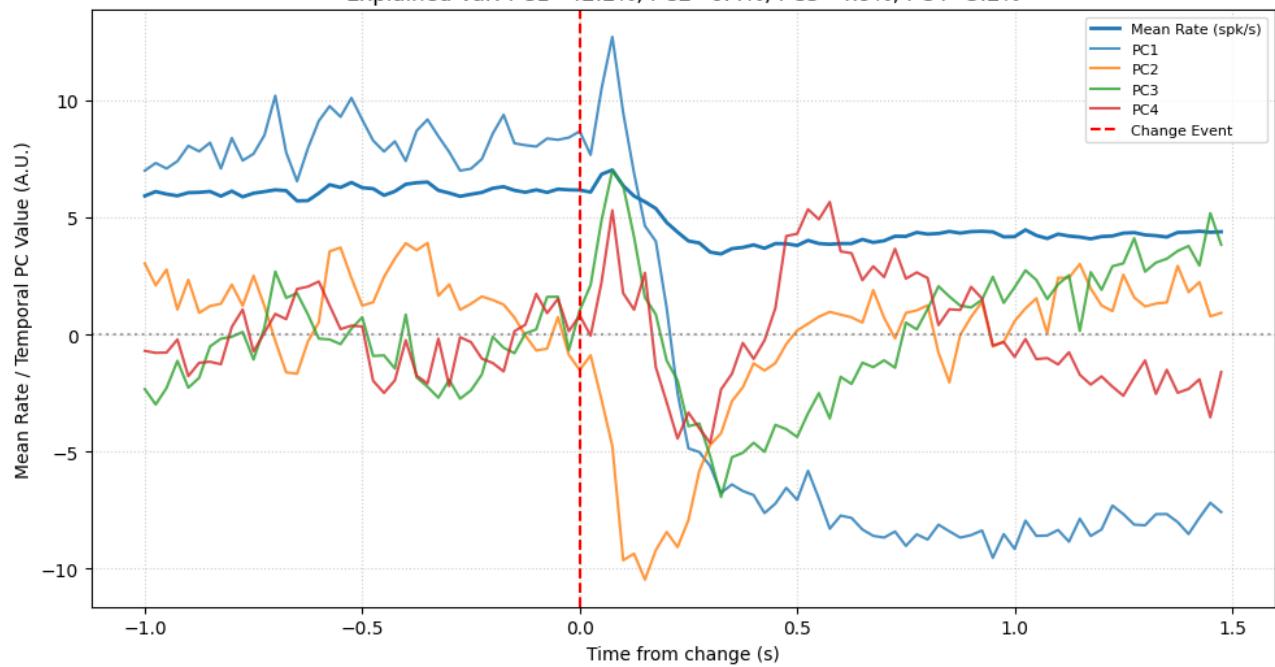
SNC | 8 units | 2,835,425 #spikes  
 Explained Var: PC1=51.0%, PC2=24.0%, PC3=10.7%, PC4=6.4%



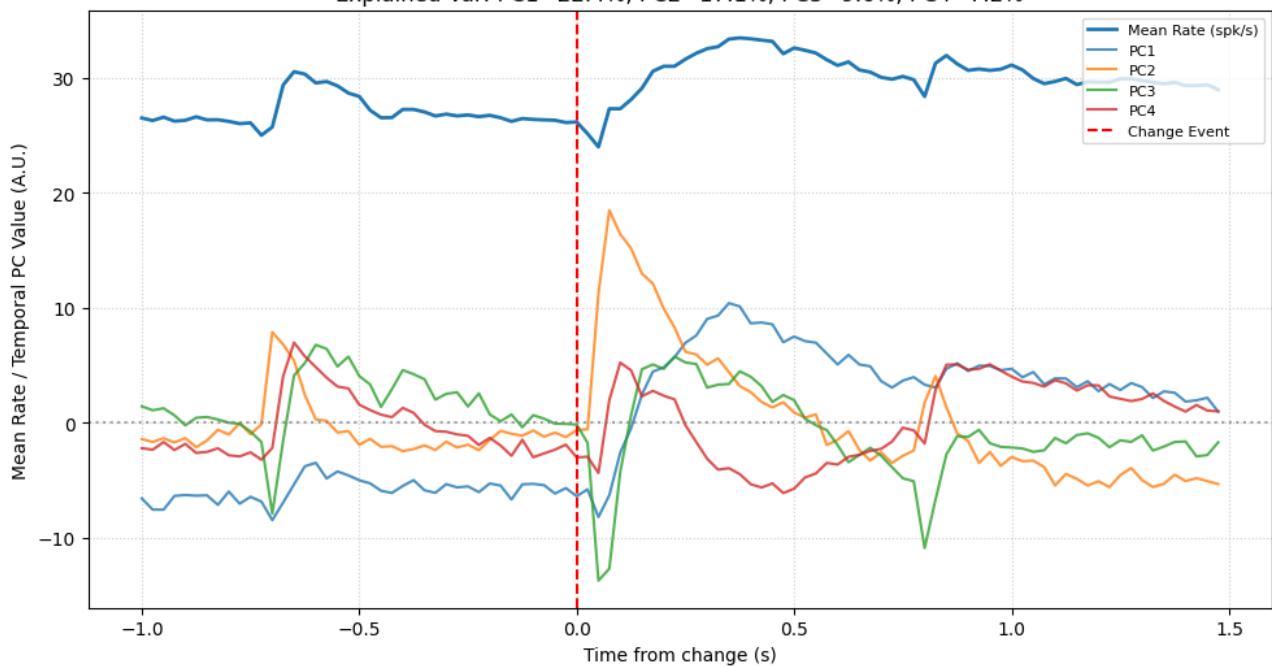
SPFp | 88 units | 7,193,848 #spikes  
 Explained Var: PC1=29.1%, PC2=14.0%, PC3=8.7%, PC4=5.3%

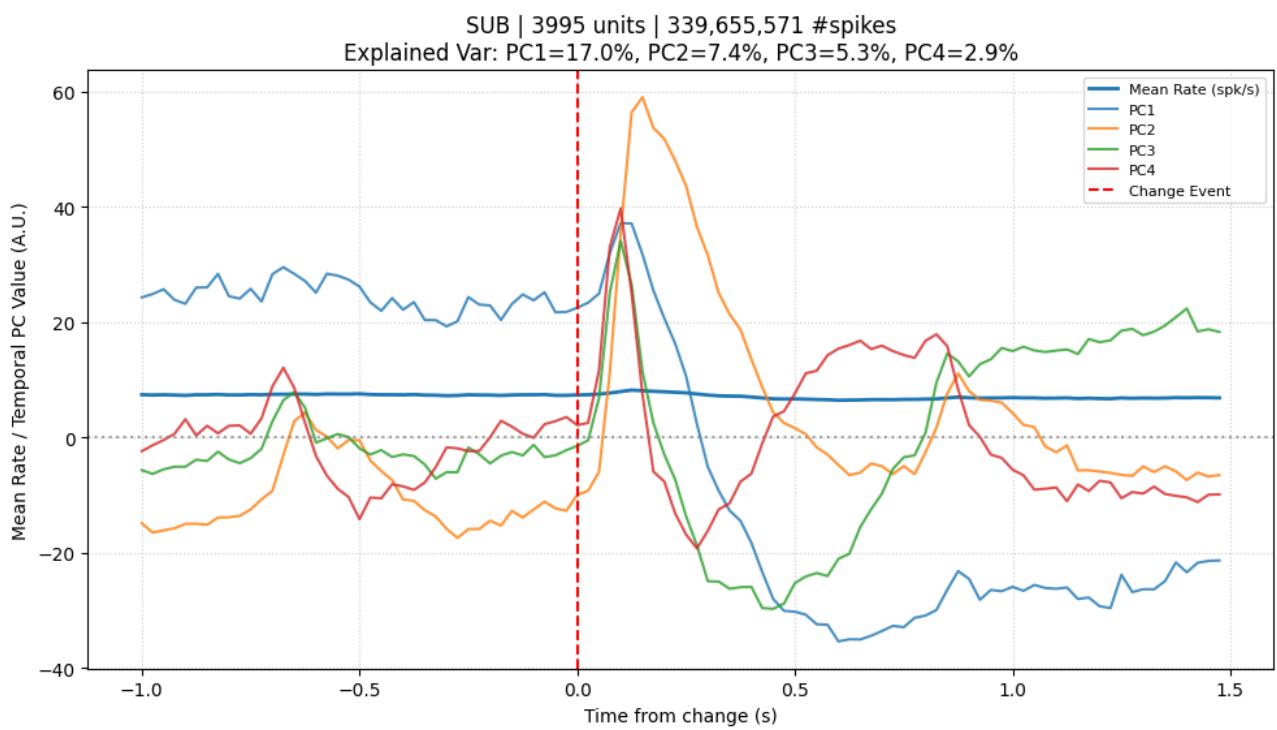


SSp | 153 units | 9,124,051 #spikes  
 Explained Var: PC1=42.2%, PC2=6.4%, PC3=4.9%, PC4=3.2%

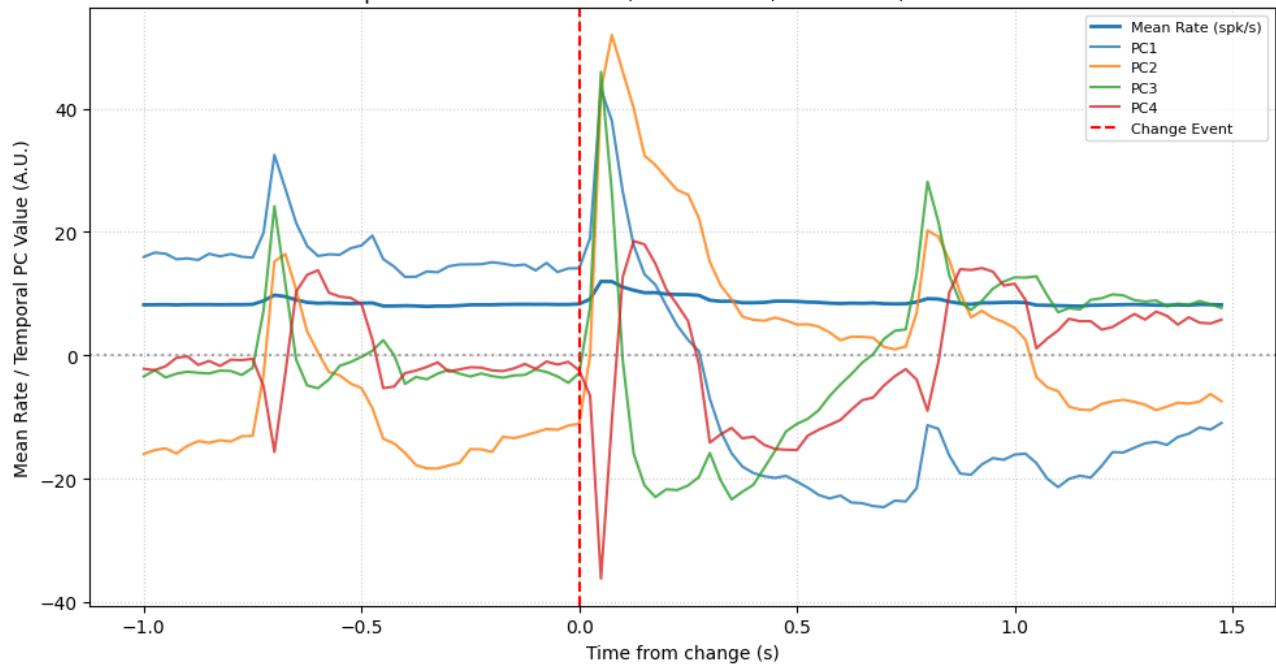


SNr | 142 units | 41,287,388 #spikes  
 Explained Var: PC1=22.4%, PC2=17.1%, PC3=9.6%, PC4=7.2%

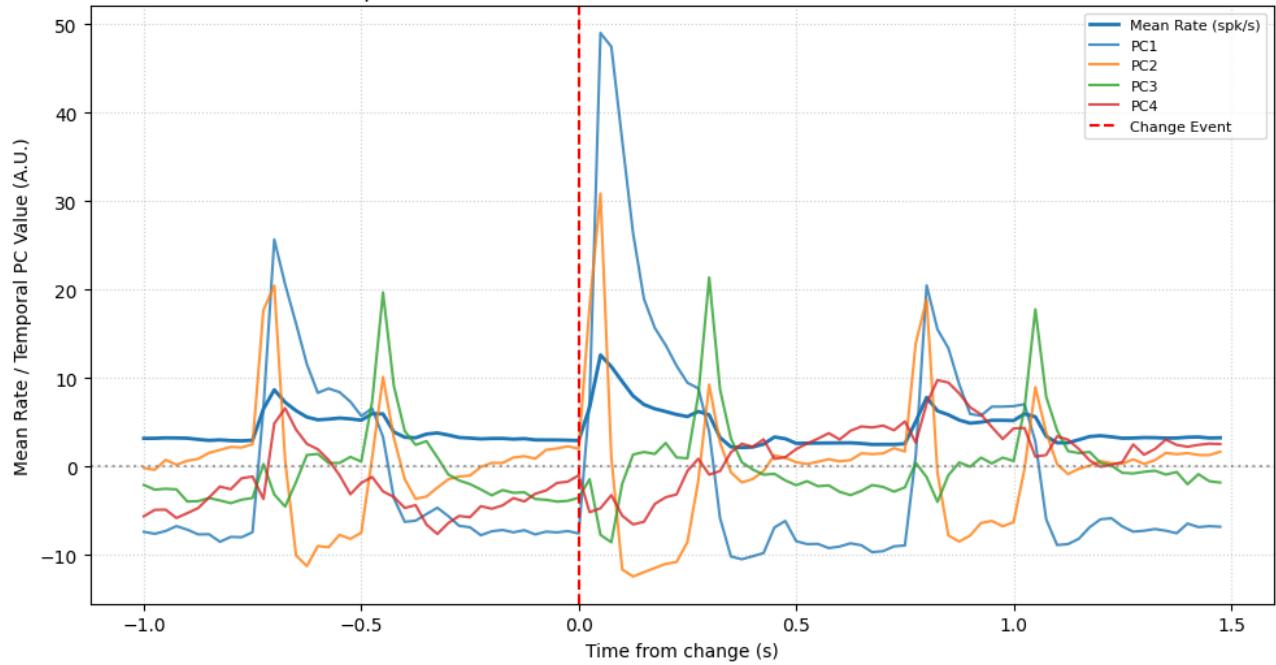




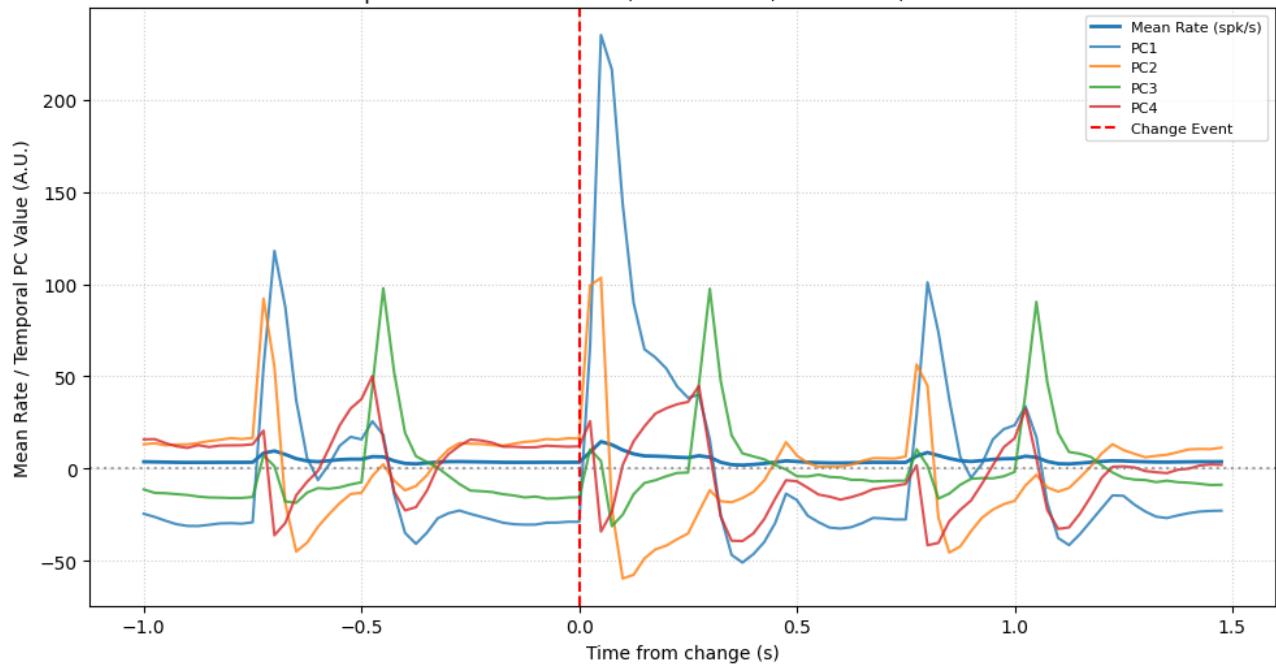
TH | 1660 units | 153,688,287 #spikes  
 Explained Var: PC1=19.8%, PC2=14.2%, PC3=8.6%, PC4=5.1%



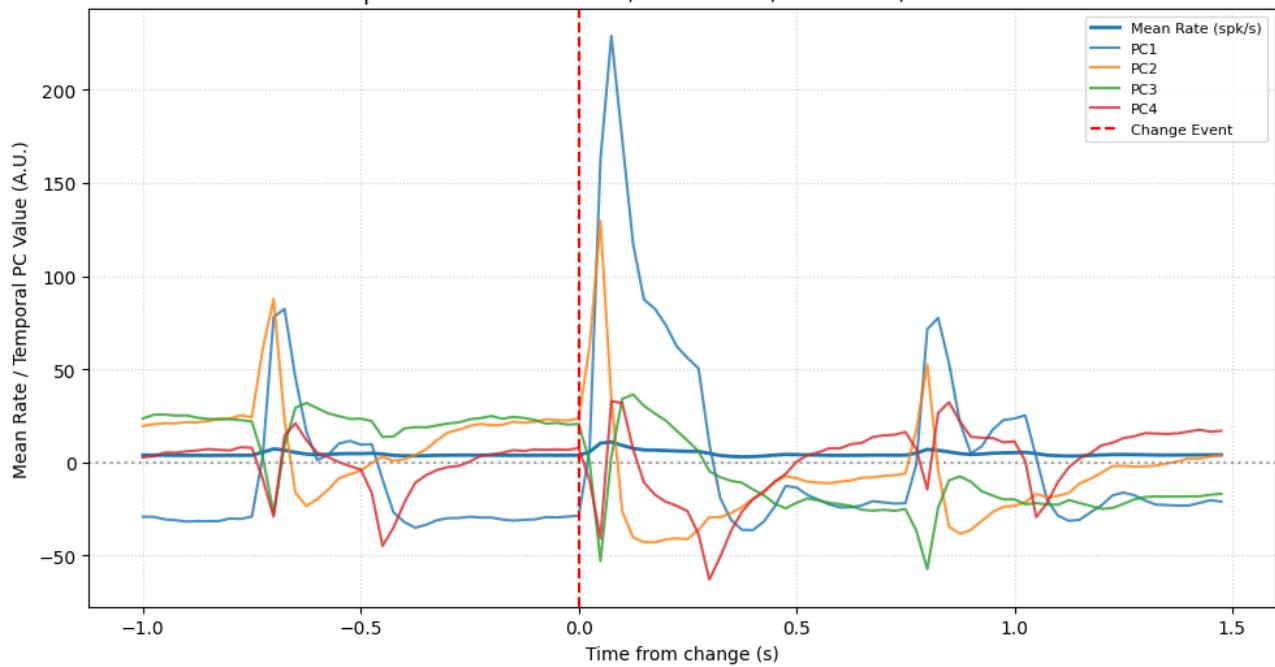
VIS | 381 units | 18,137,345 #spikes  
 Explained Var: PC1=38.7%, PC2=12.7%, PC3=5.7%, PC4=4.3%

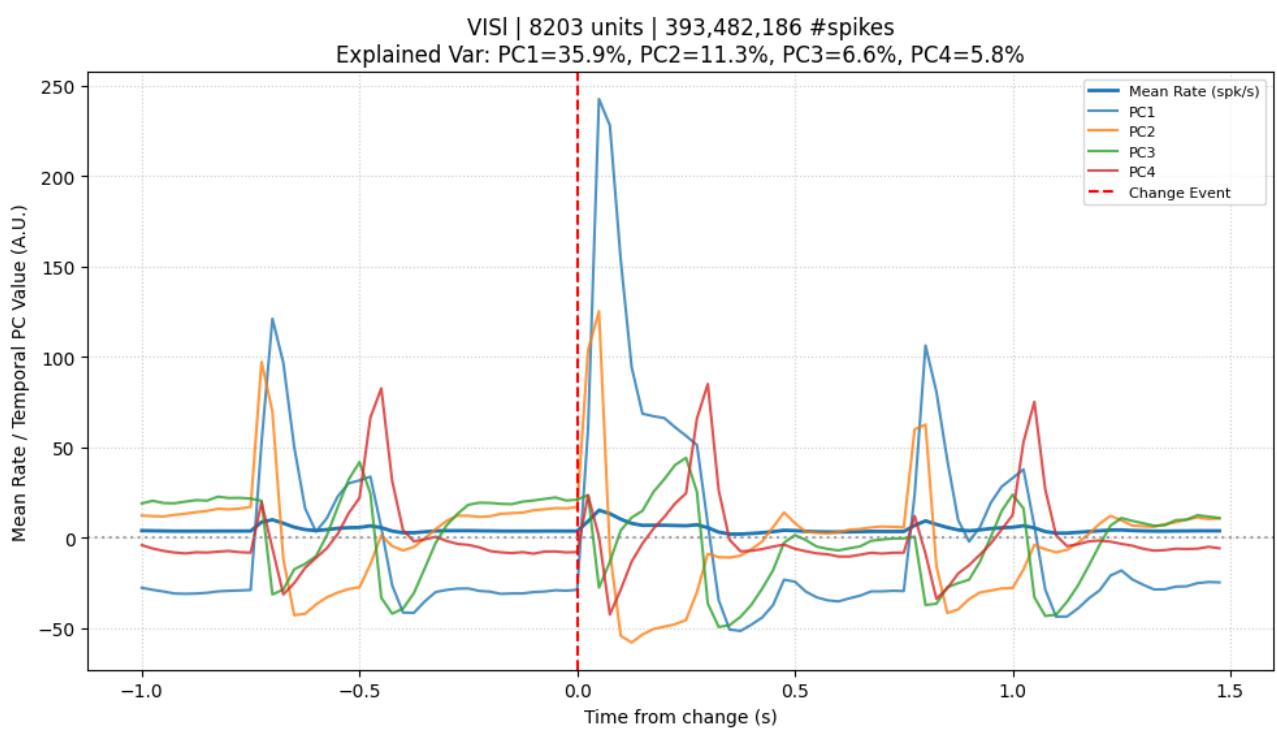


VISal | 7441 units | 382,868,073 #spikes  
 Explained Var: PC1=34.4%, PC2=10.3%, PC3=6.9%, PC4=5.6%

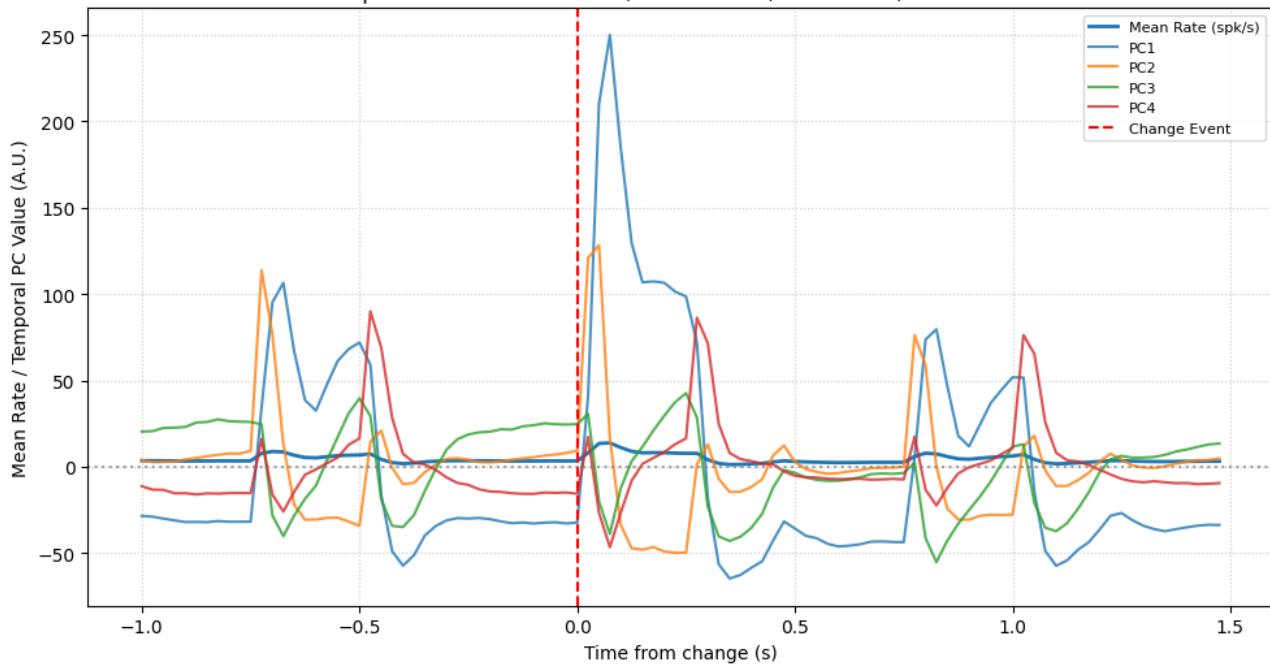


VISam | 7398 units | 341,591,849 #spikes  
 Explained Var: PC1=31.4%, PC2=10.3%, PC3=7.1%, PC4=4.4%

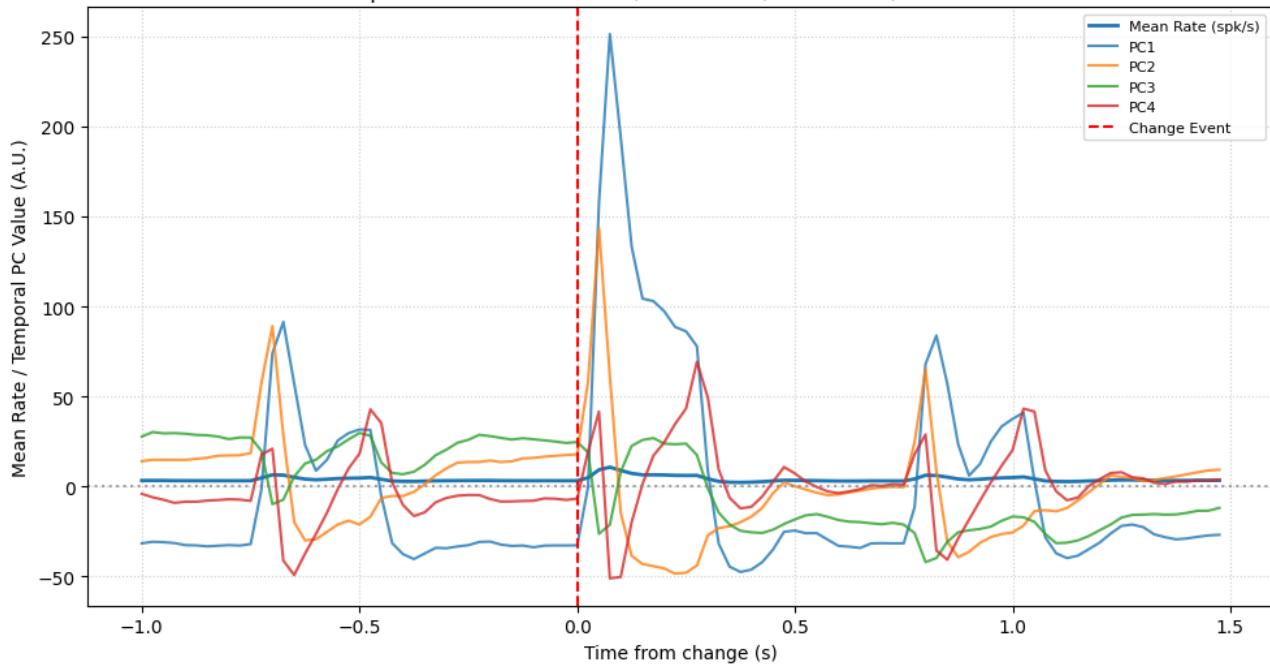




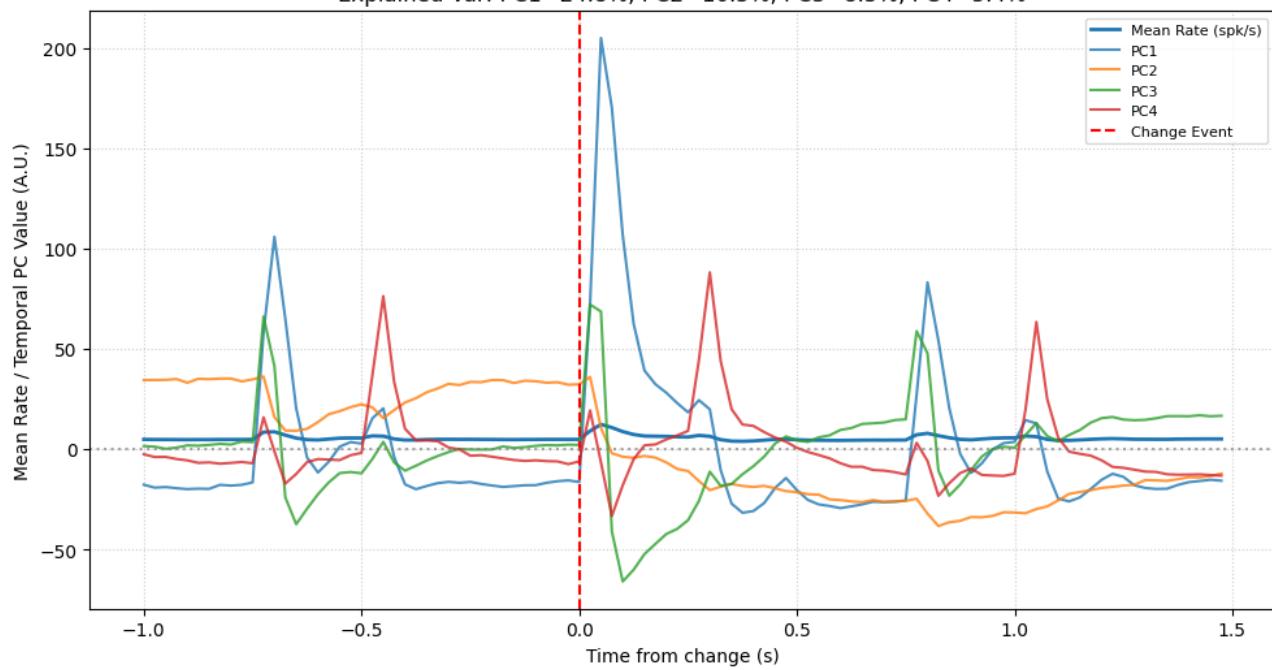
VISp | 8818 units | 441,276,196 #spikes  
 Explained Var: PC1=44.8%, PC2=10.3%, PC3=6.3%, PC4=5.9%



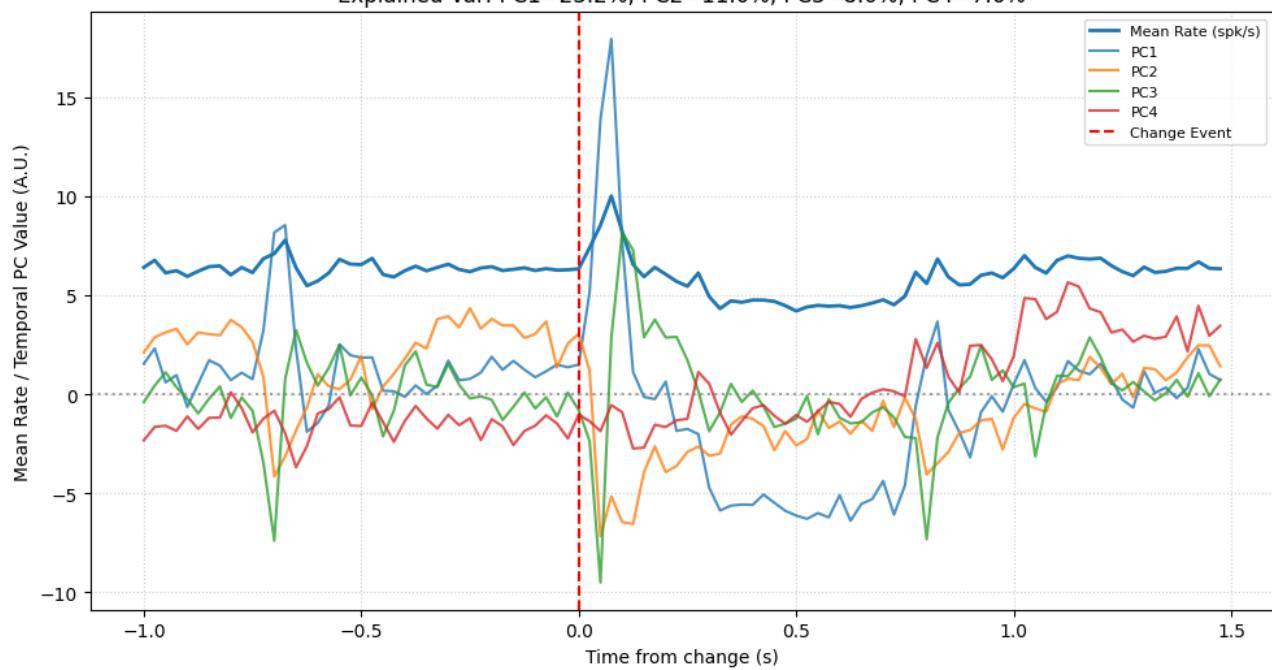
VISpm | 8885 units | 372,855,455 #spikes  
 Explained Var: PC1=34.5%, PC2=9.1%, PC3=5.8%, PC4=4.7%



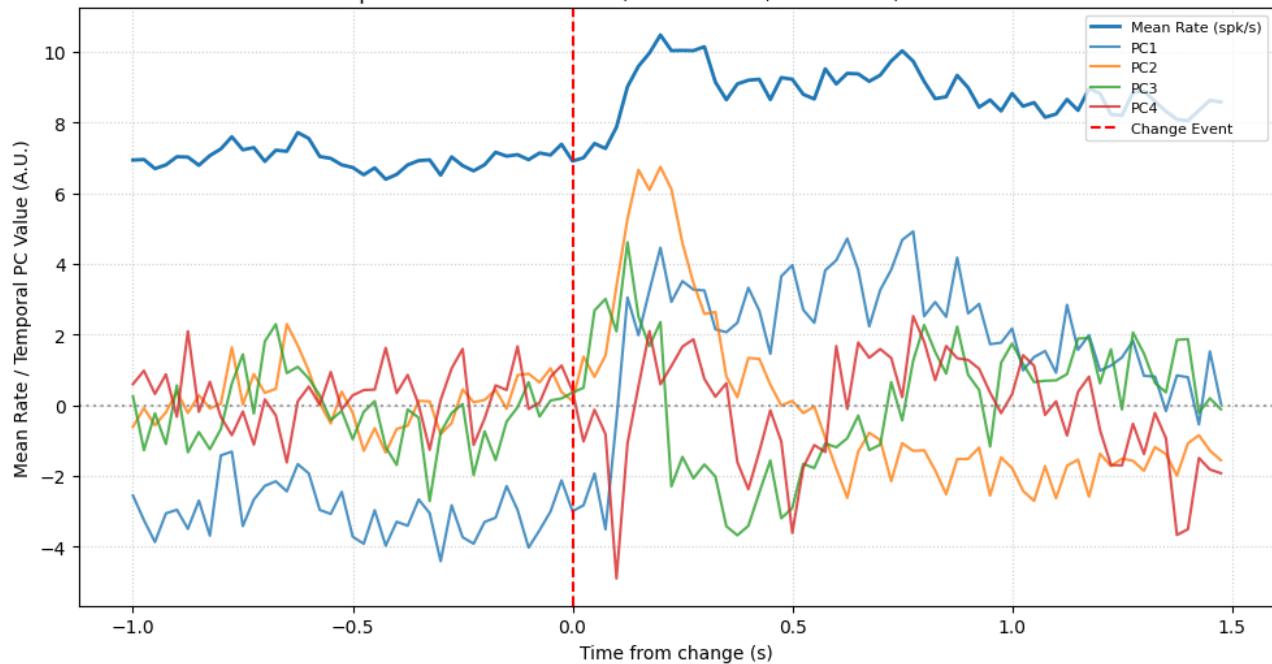
VISrl | 6310 units | 361,098,303 #spikes  
 Explained Var: PC1=24.8%, PC2=10.5%, PC3=8.5%, PC4=5.4%

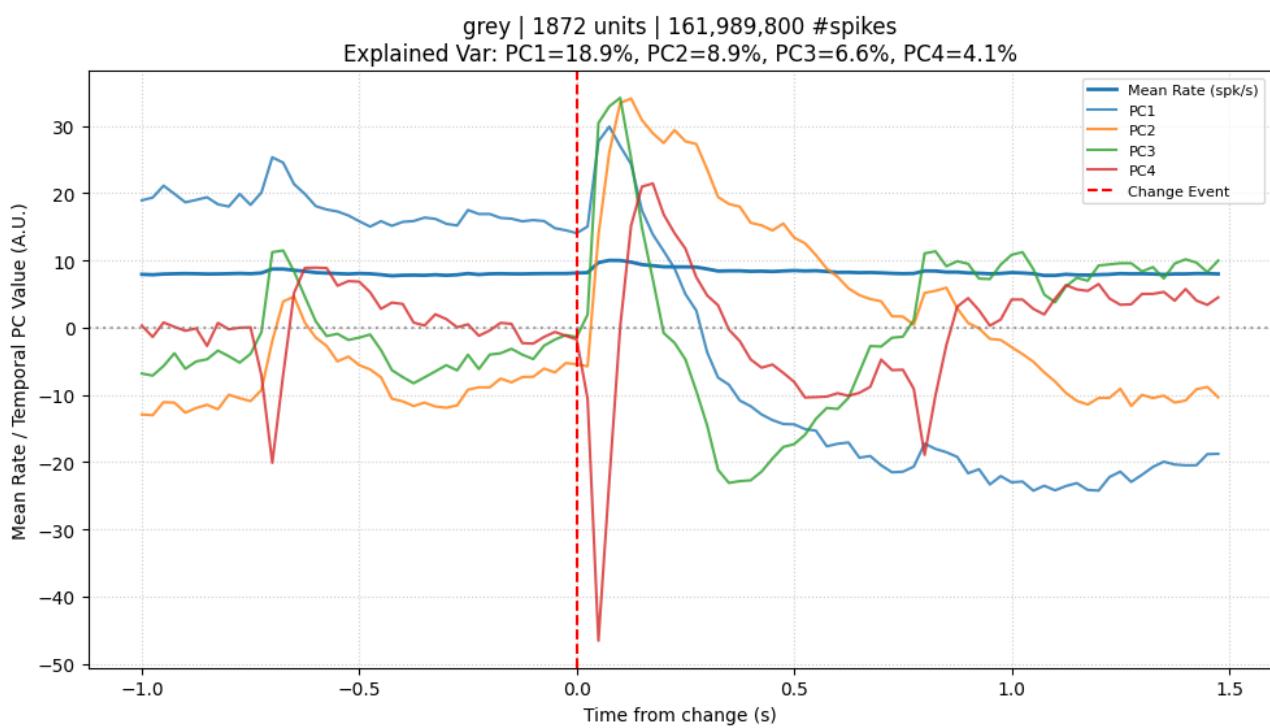
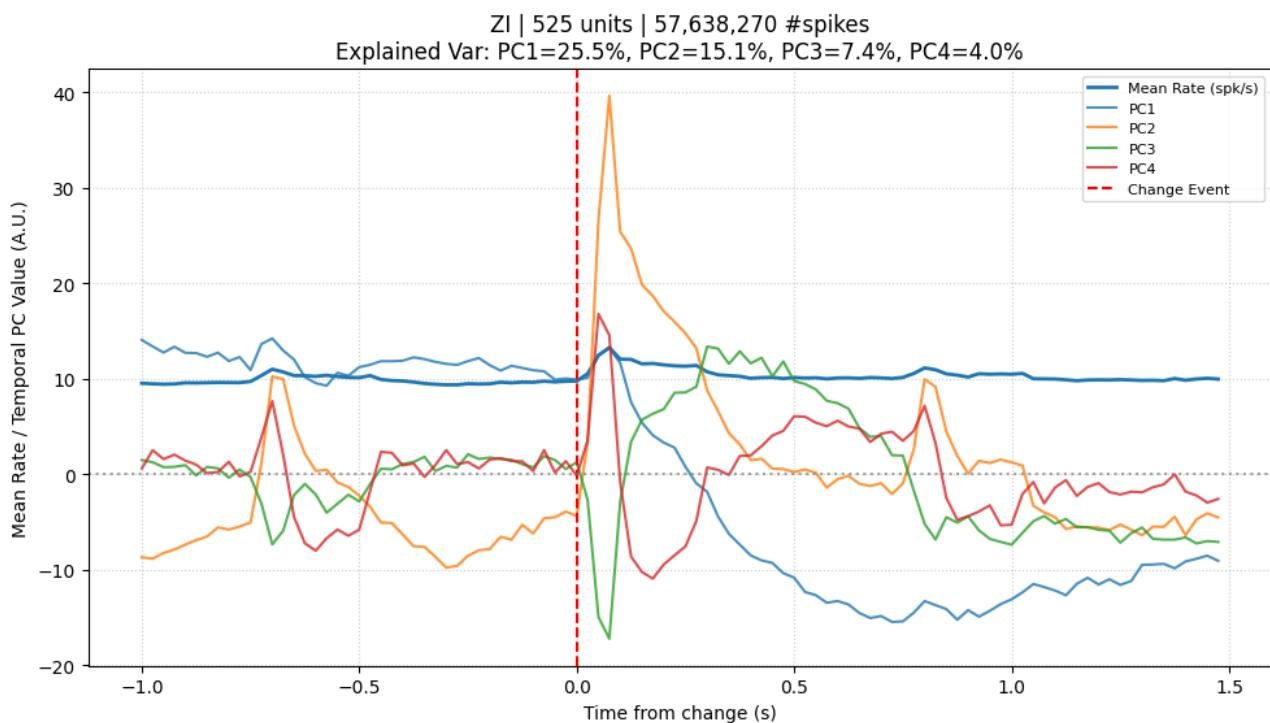


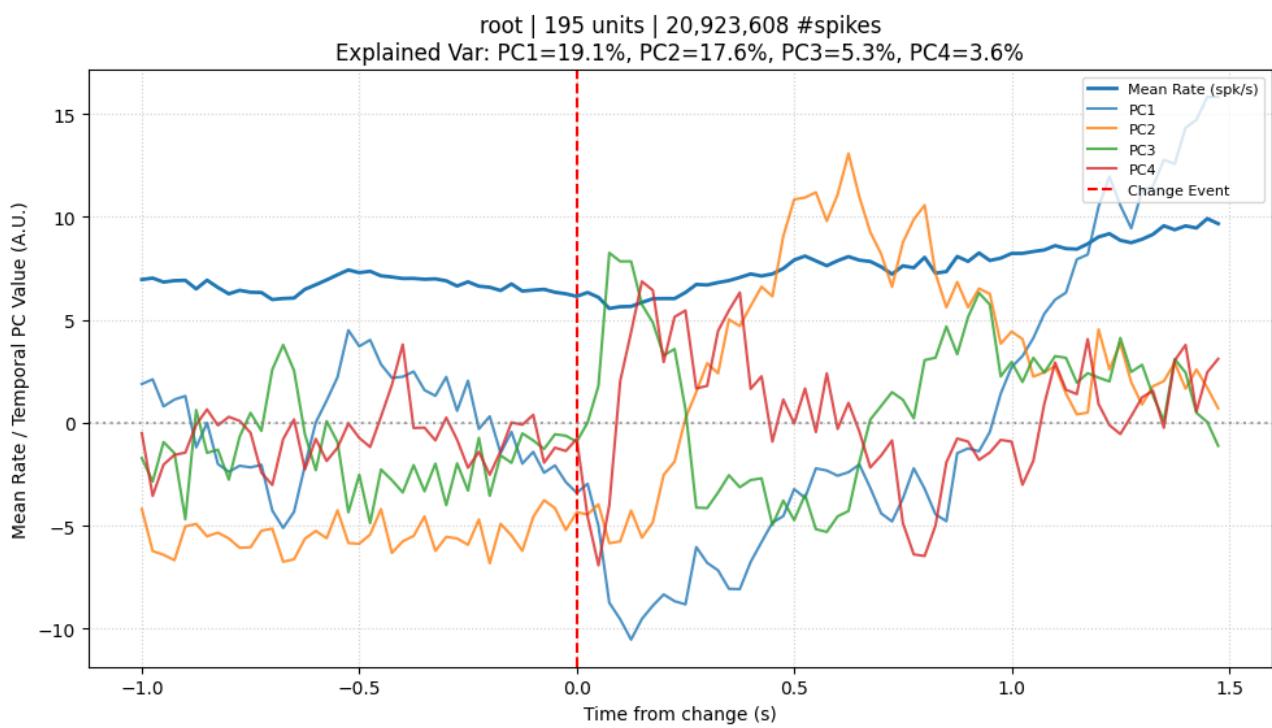
VISrll | 64 units | 4,015,008 #spikes  
 Explained Var: PC1=23.2%, PC2=11.0%, PC3=8.0%, PC4=7.6%



VPLpc | 31 units | 2,344,822 #spikes  
Explained Var: PC1=26.4%, PC2=12.8%, PC3=7.9%, PC4=5.8%

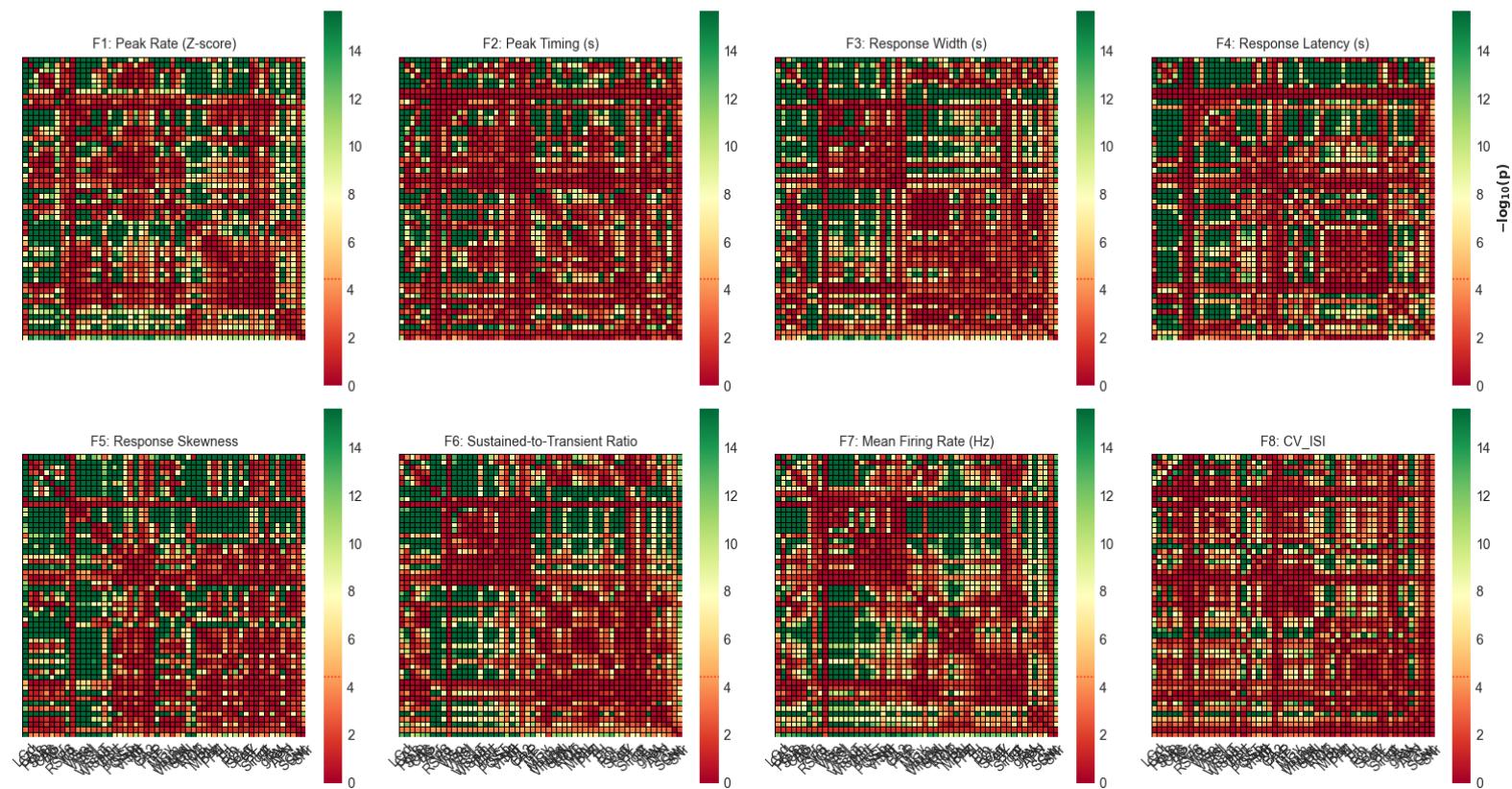




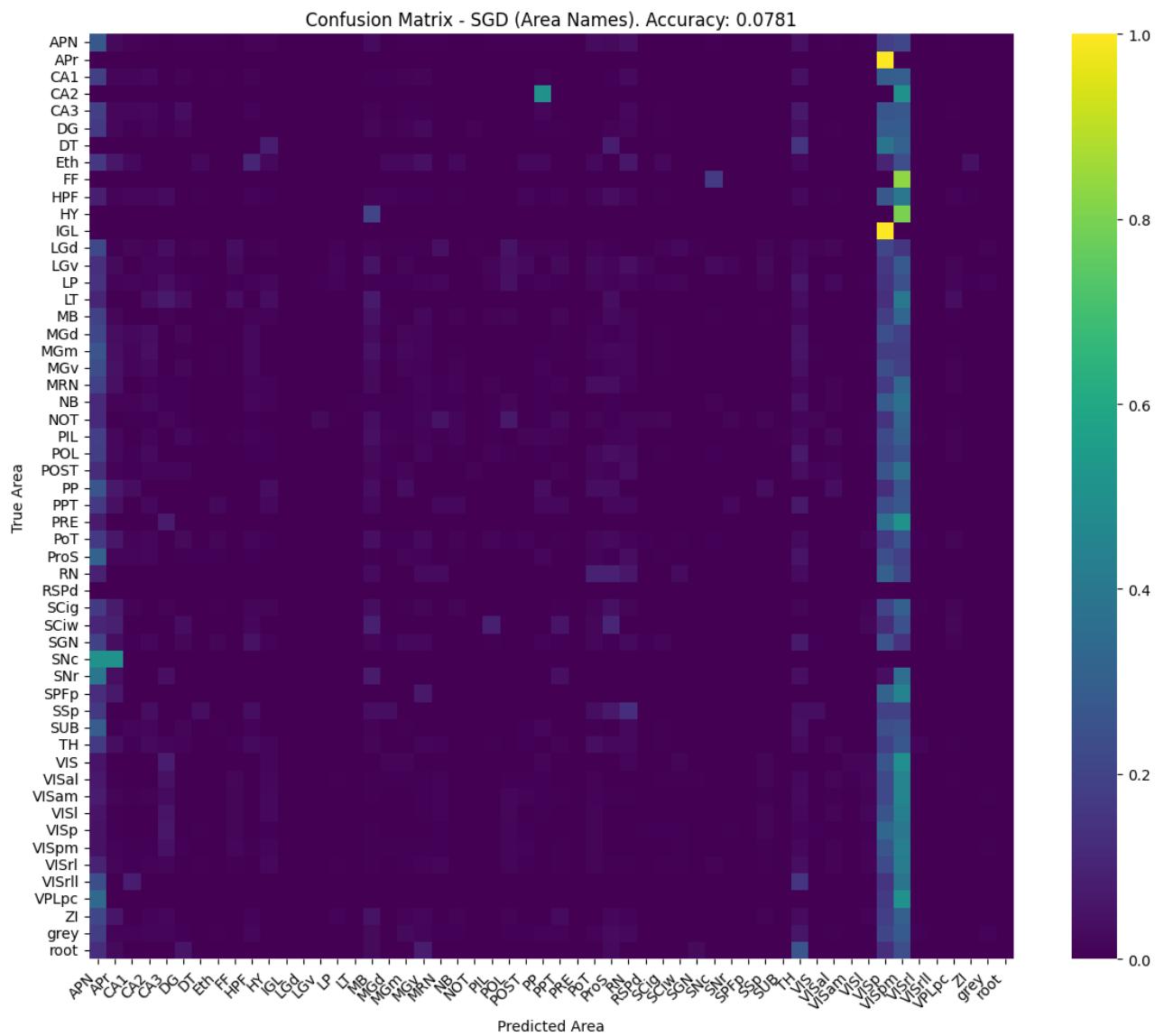


# Features dissimilarity matrix

Individual Feature Dissimilarity Matrices  $\sum -\log_{10}(p)$   
All maps clustered consistently by Omnibus dissimilarity. Significance Reference: 4.46

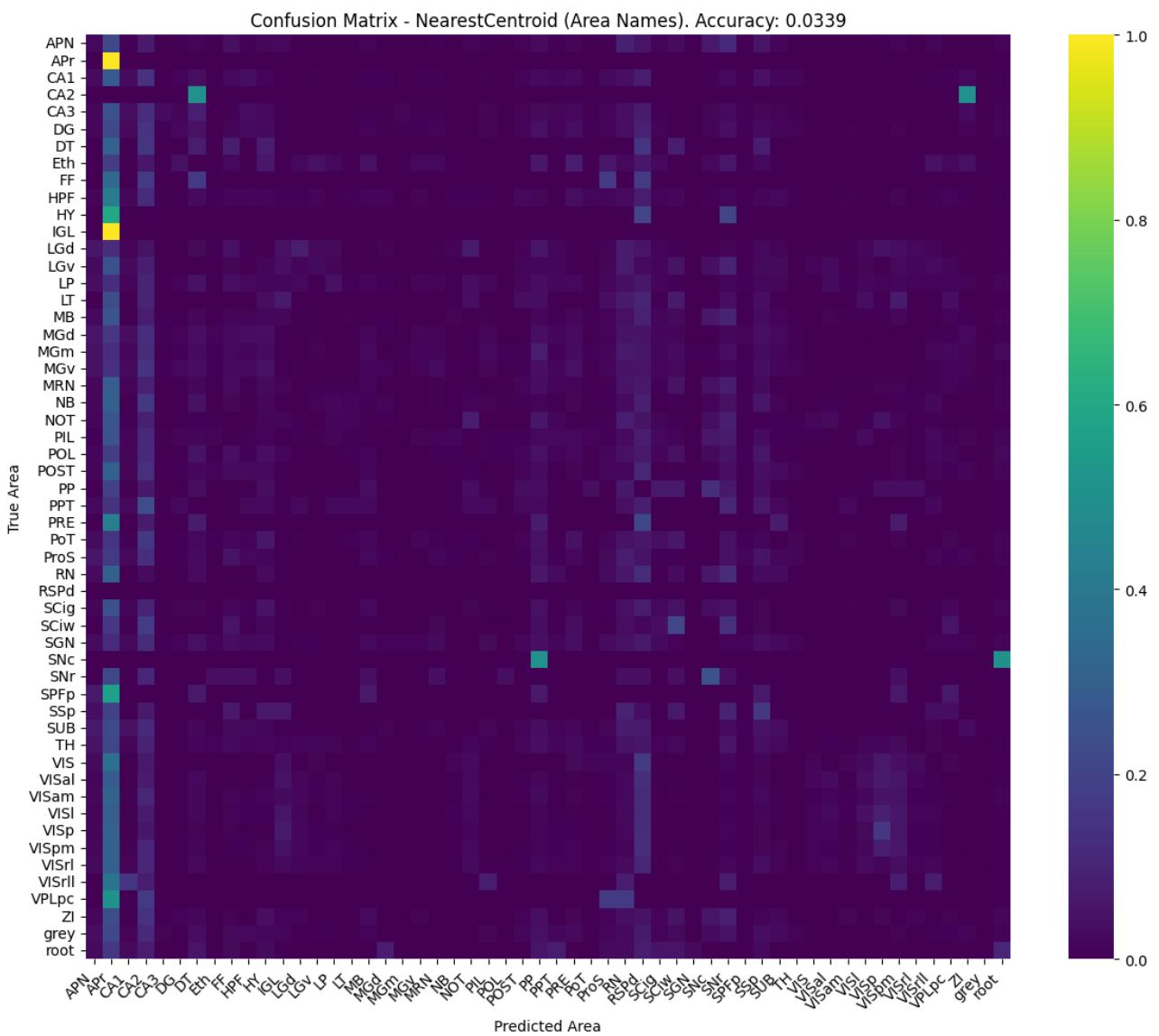


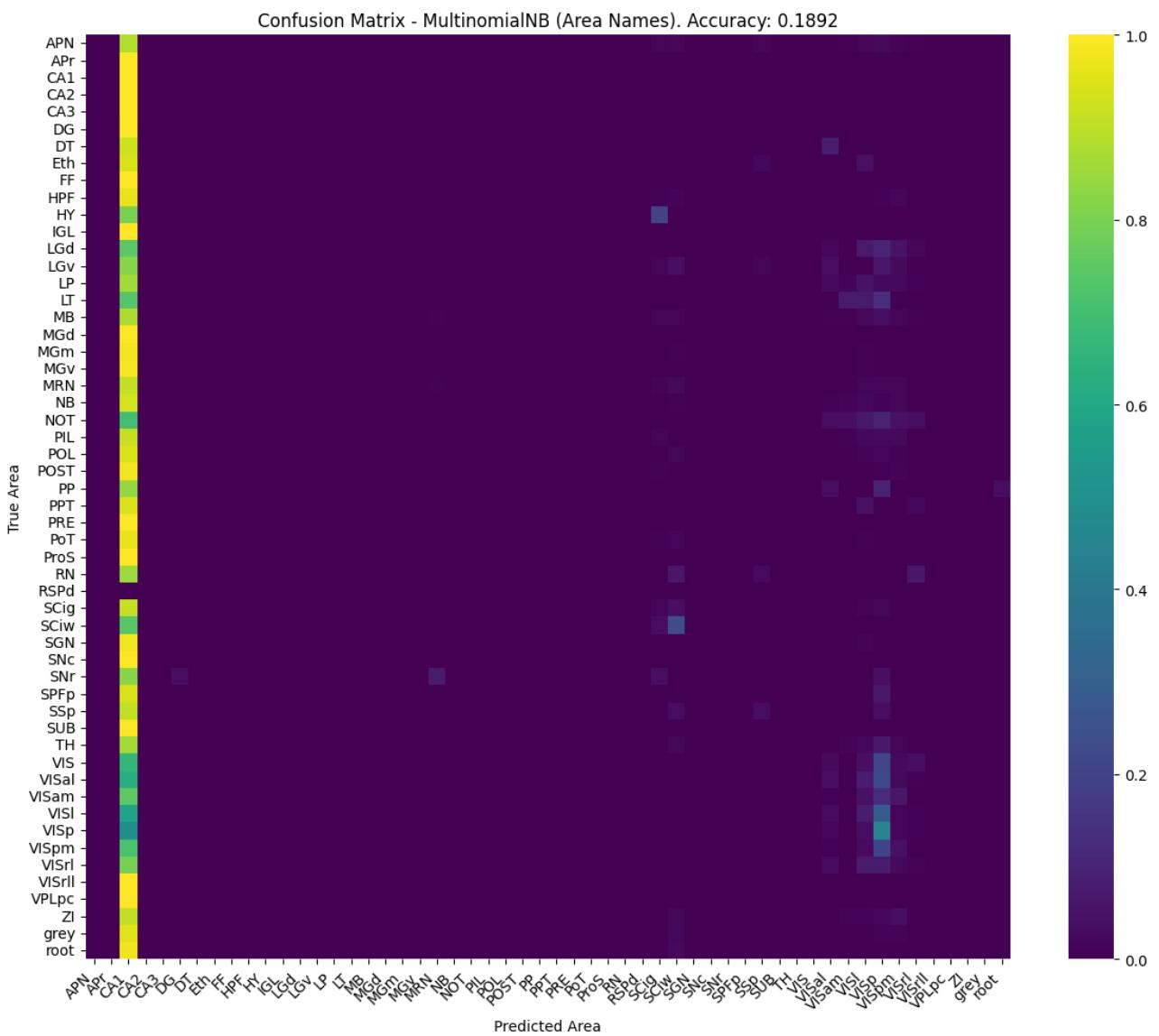
# Confusion matrices classic classifiers:





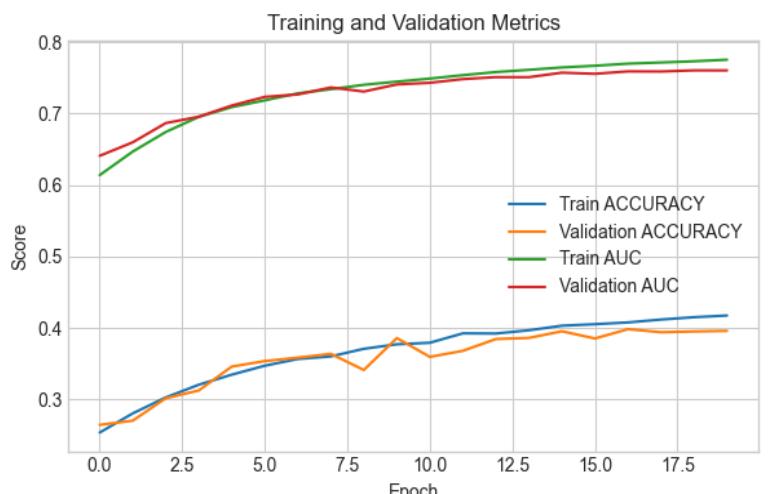
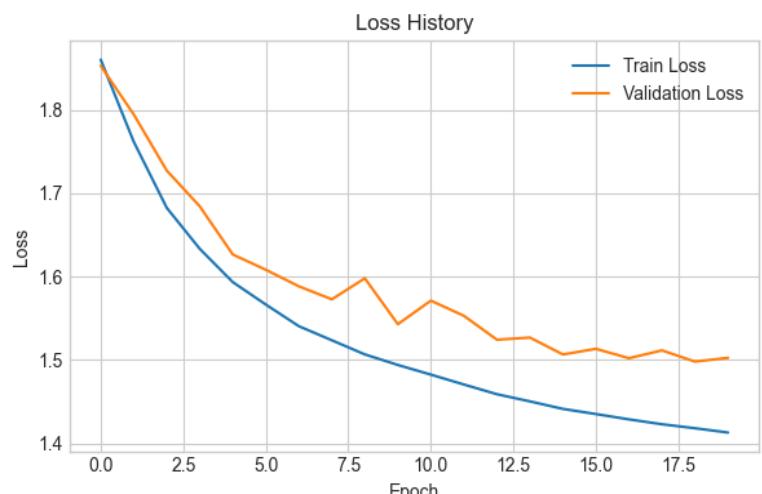




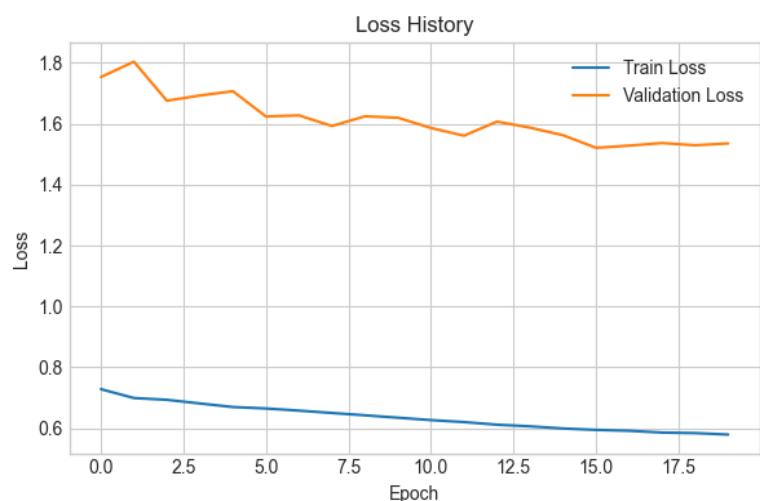
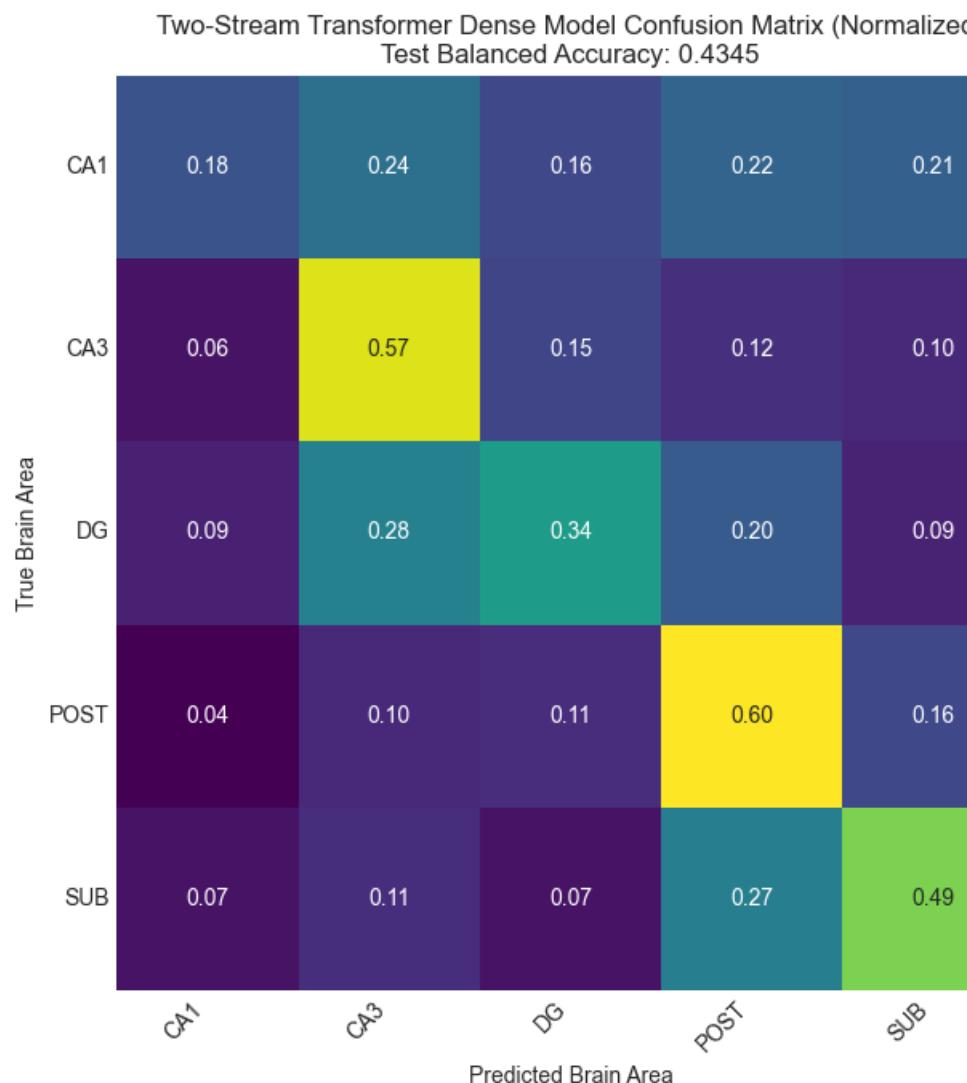


# Confusion matrices and training plots

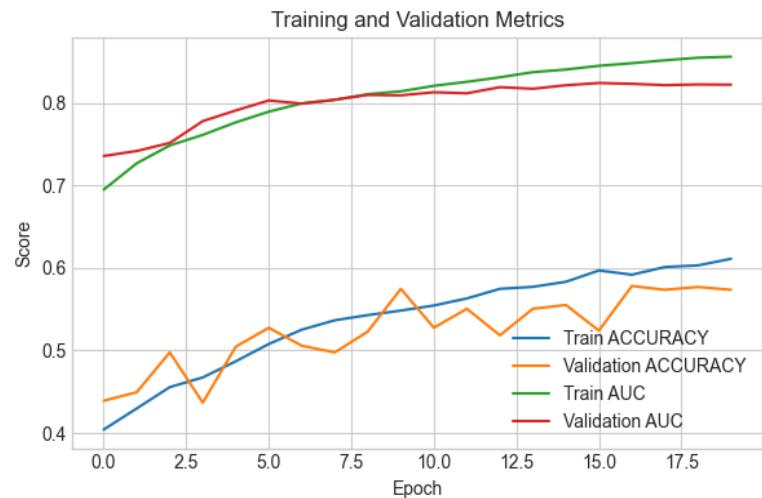
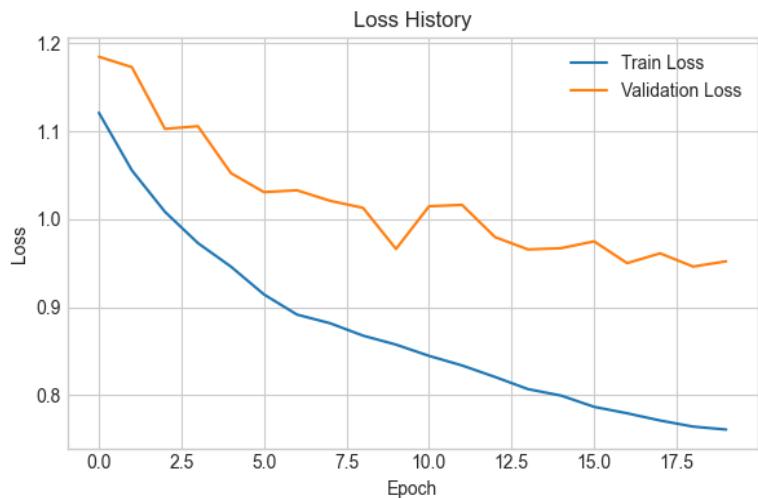
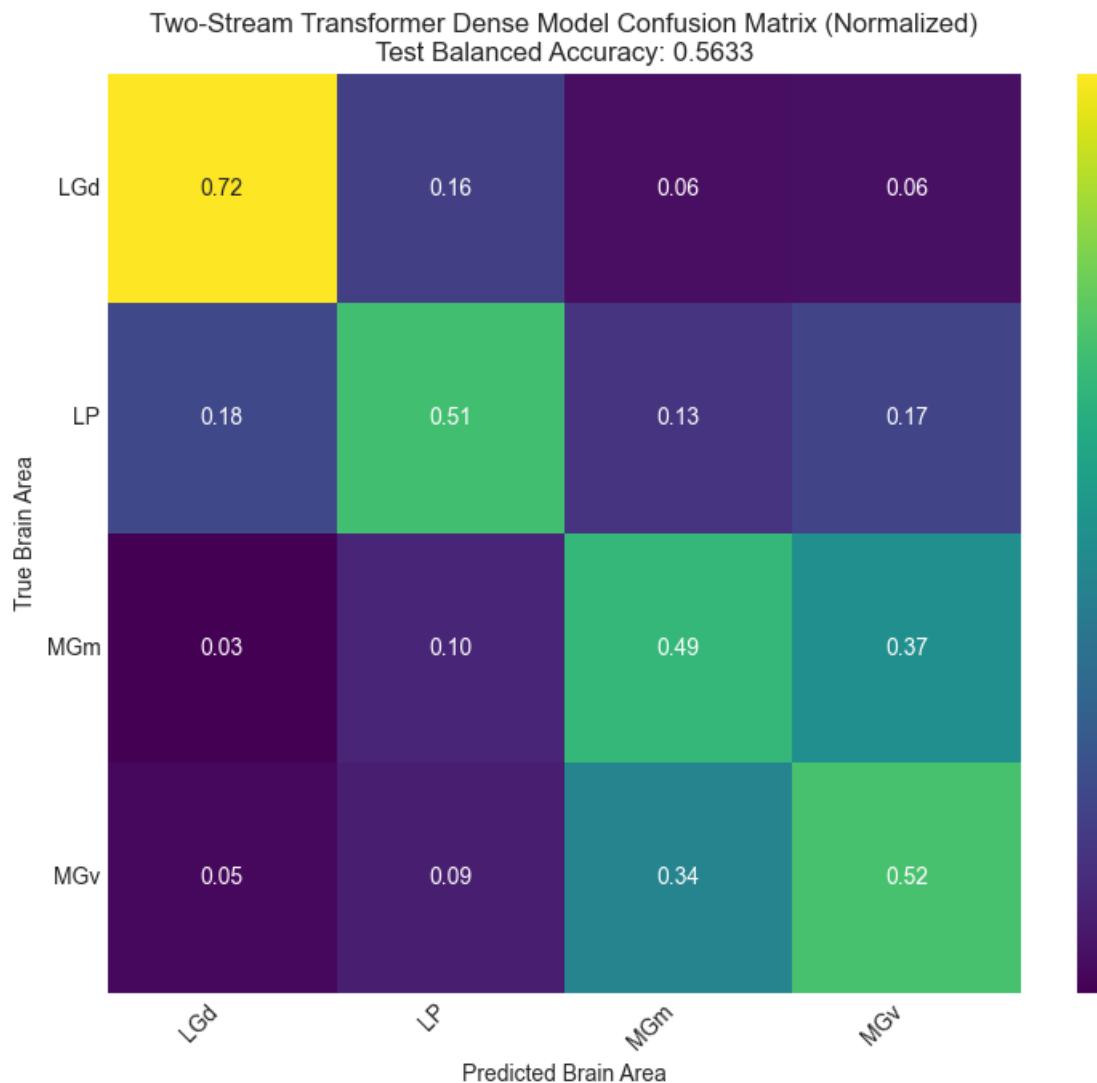
## Transformer visual



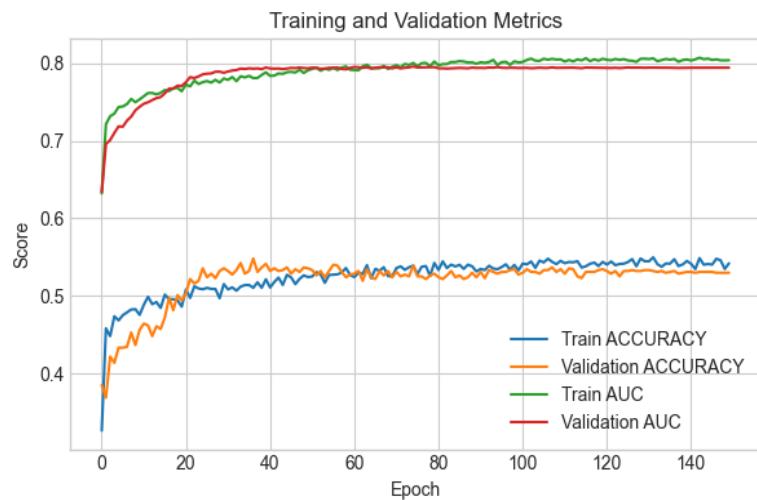
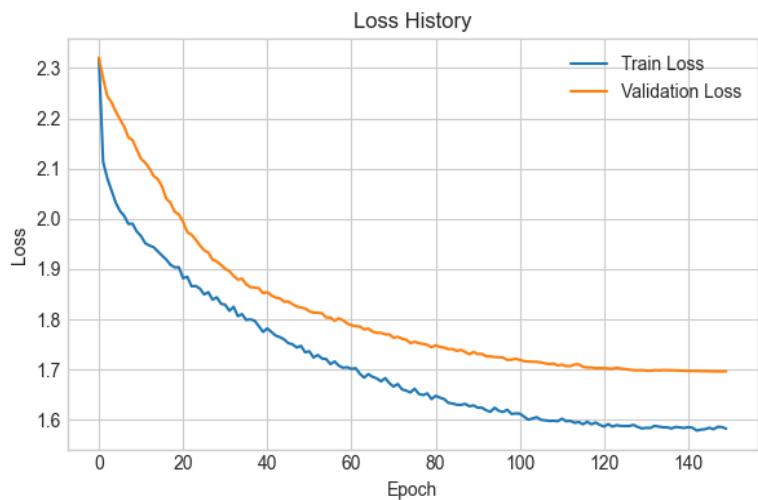
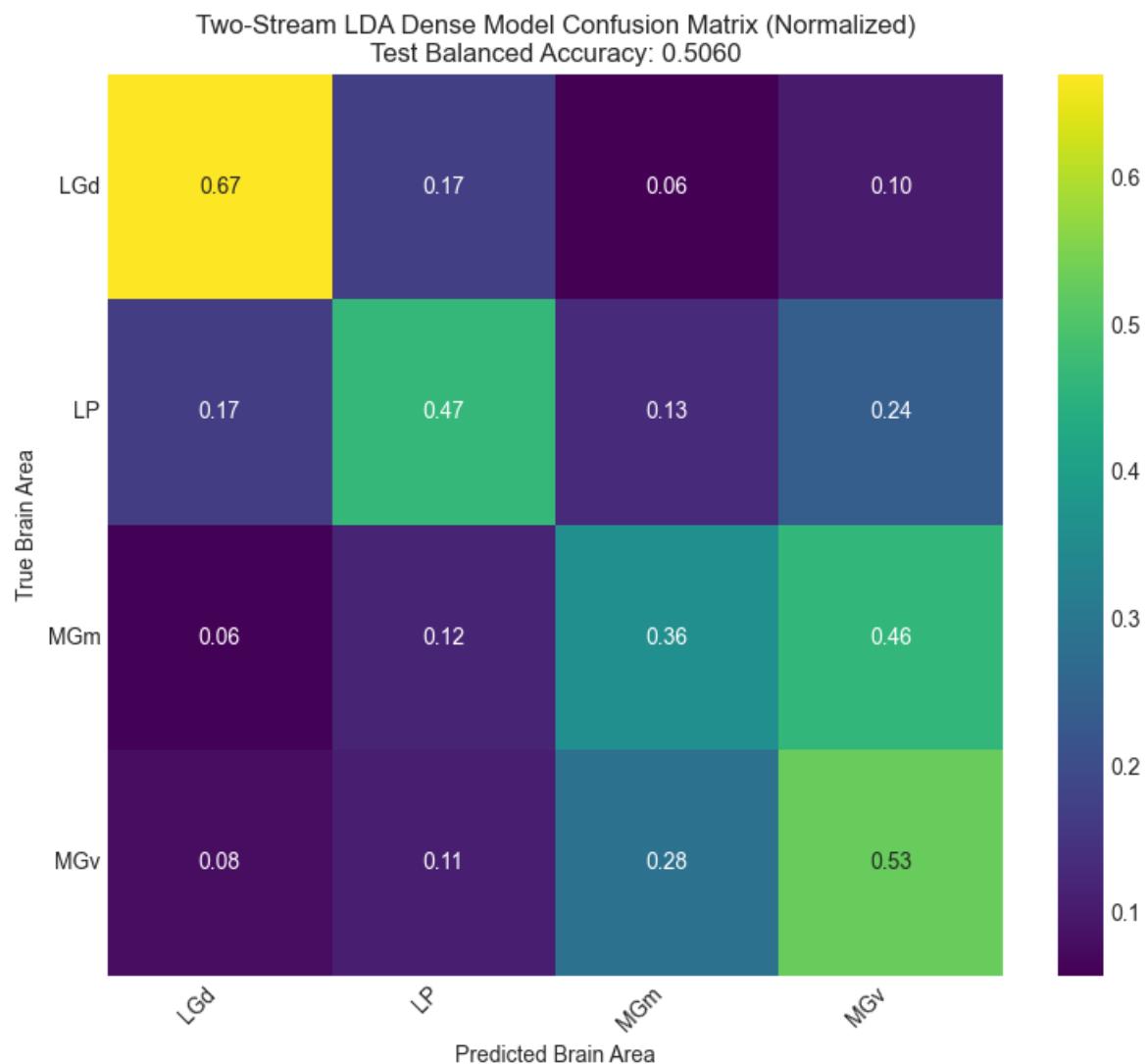
# Transformer hippocampus



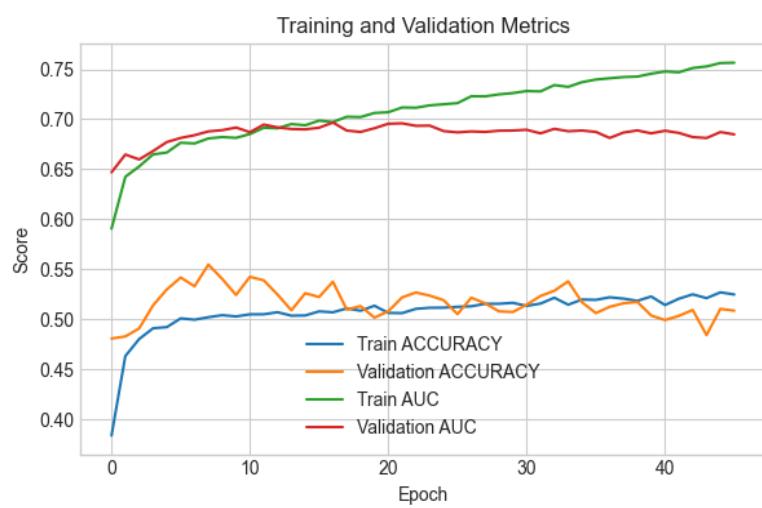
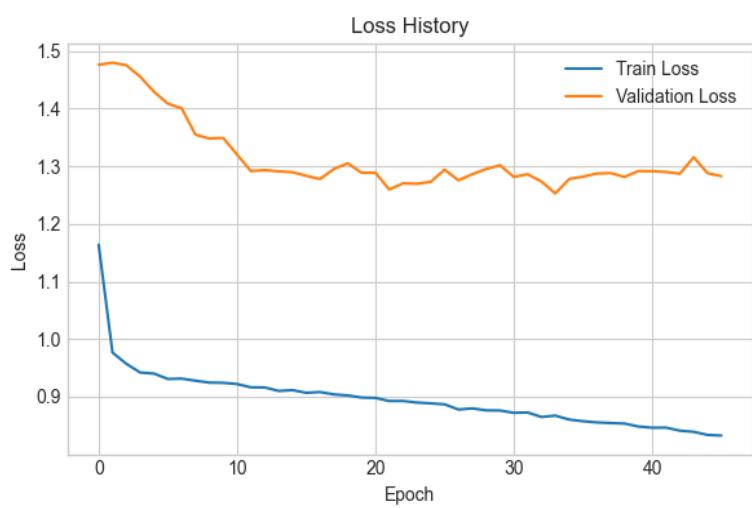
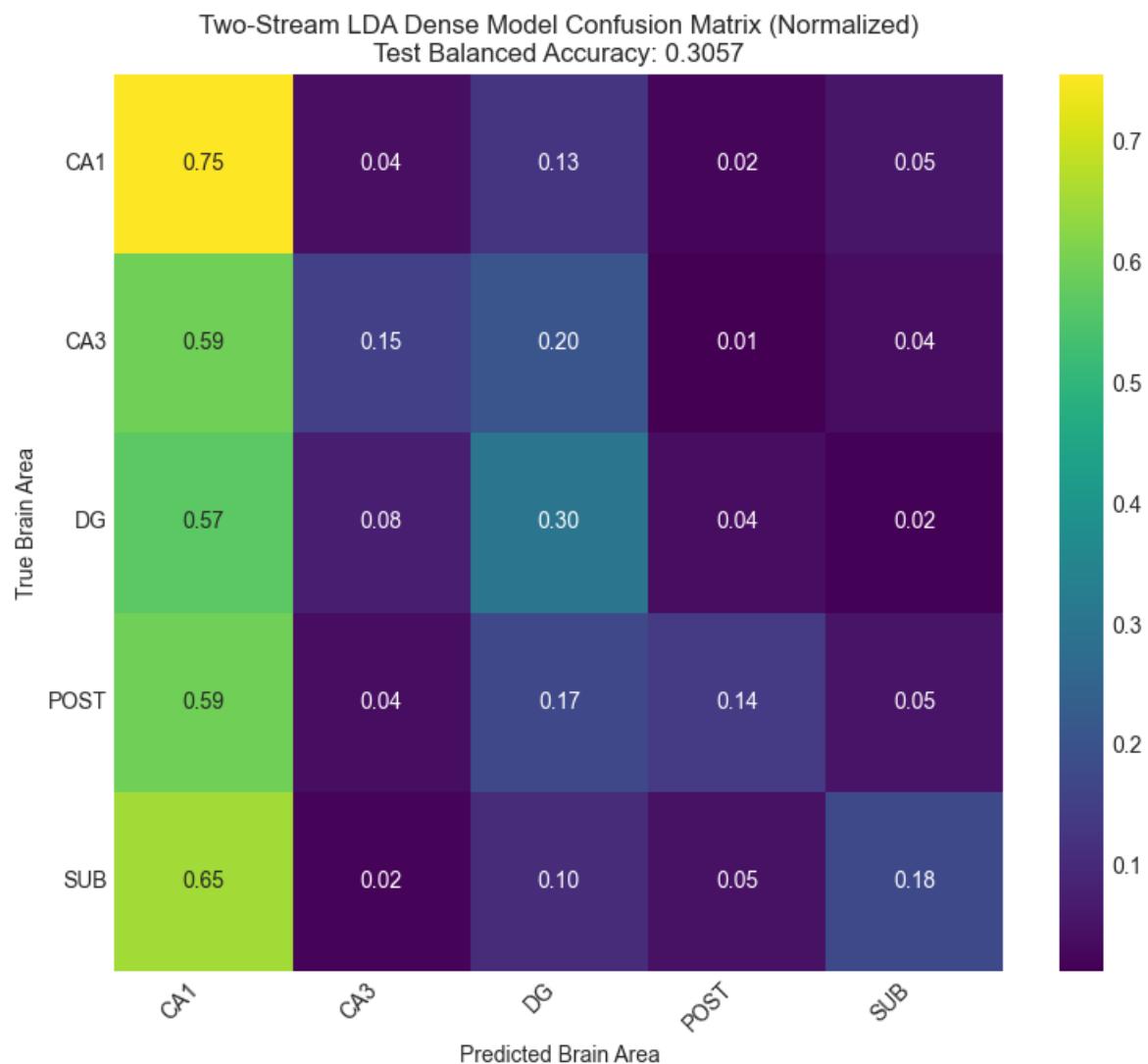
# Transformer thalamus



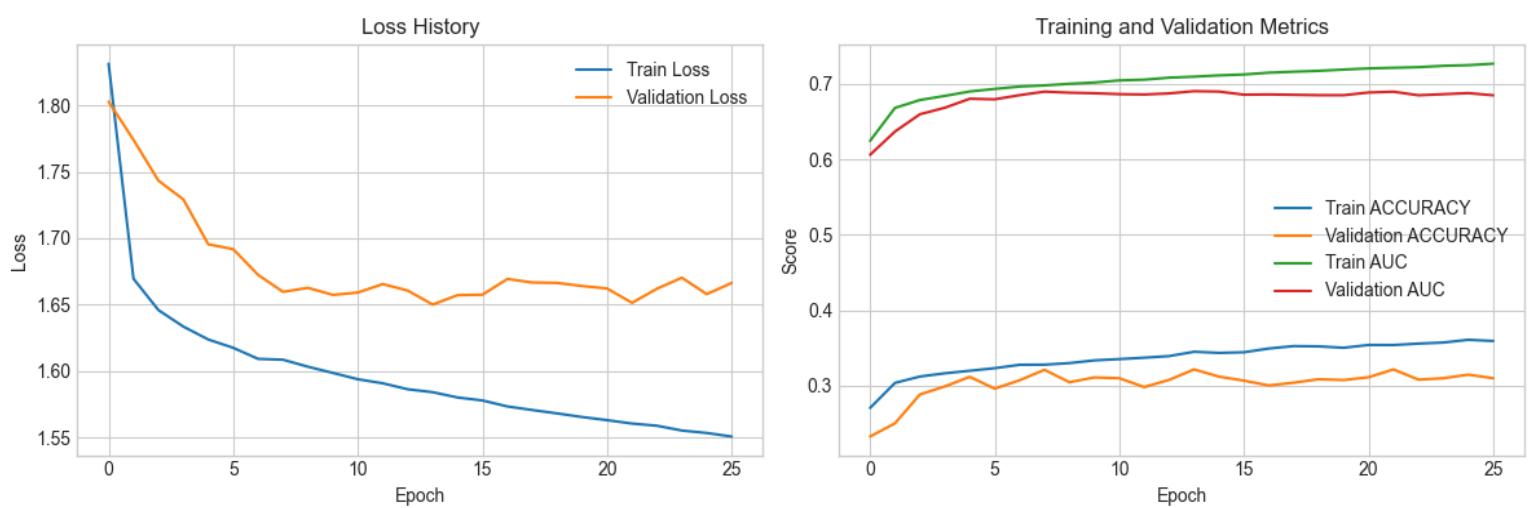
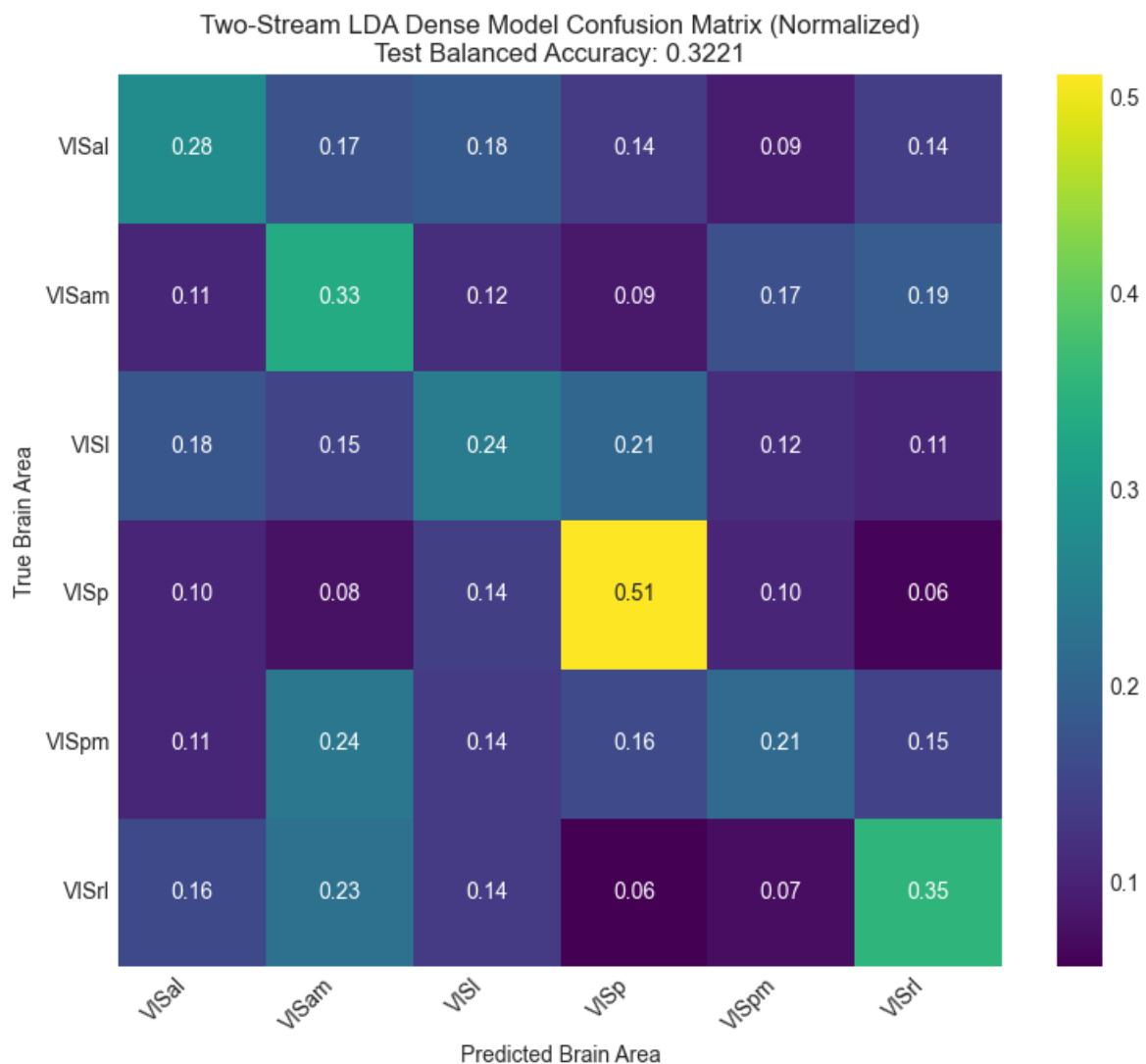
# LDA dense thalamus



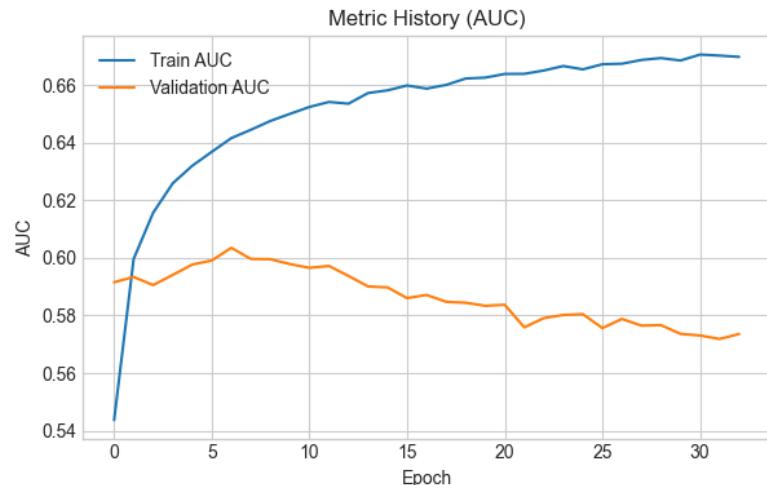
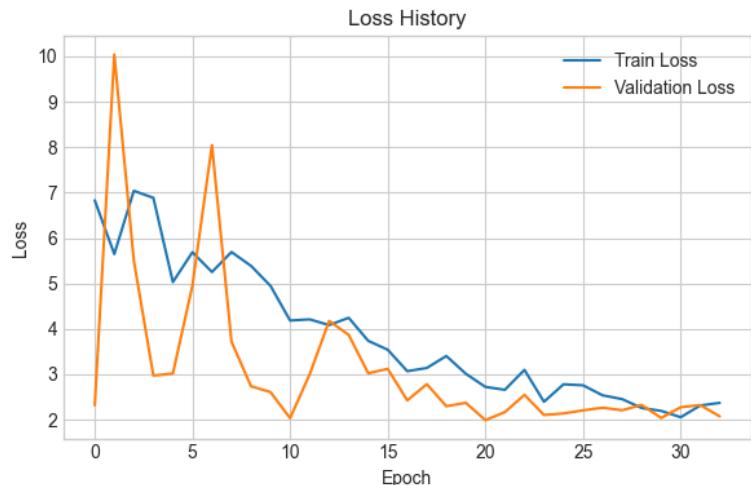
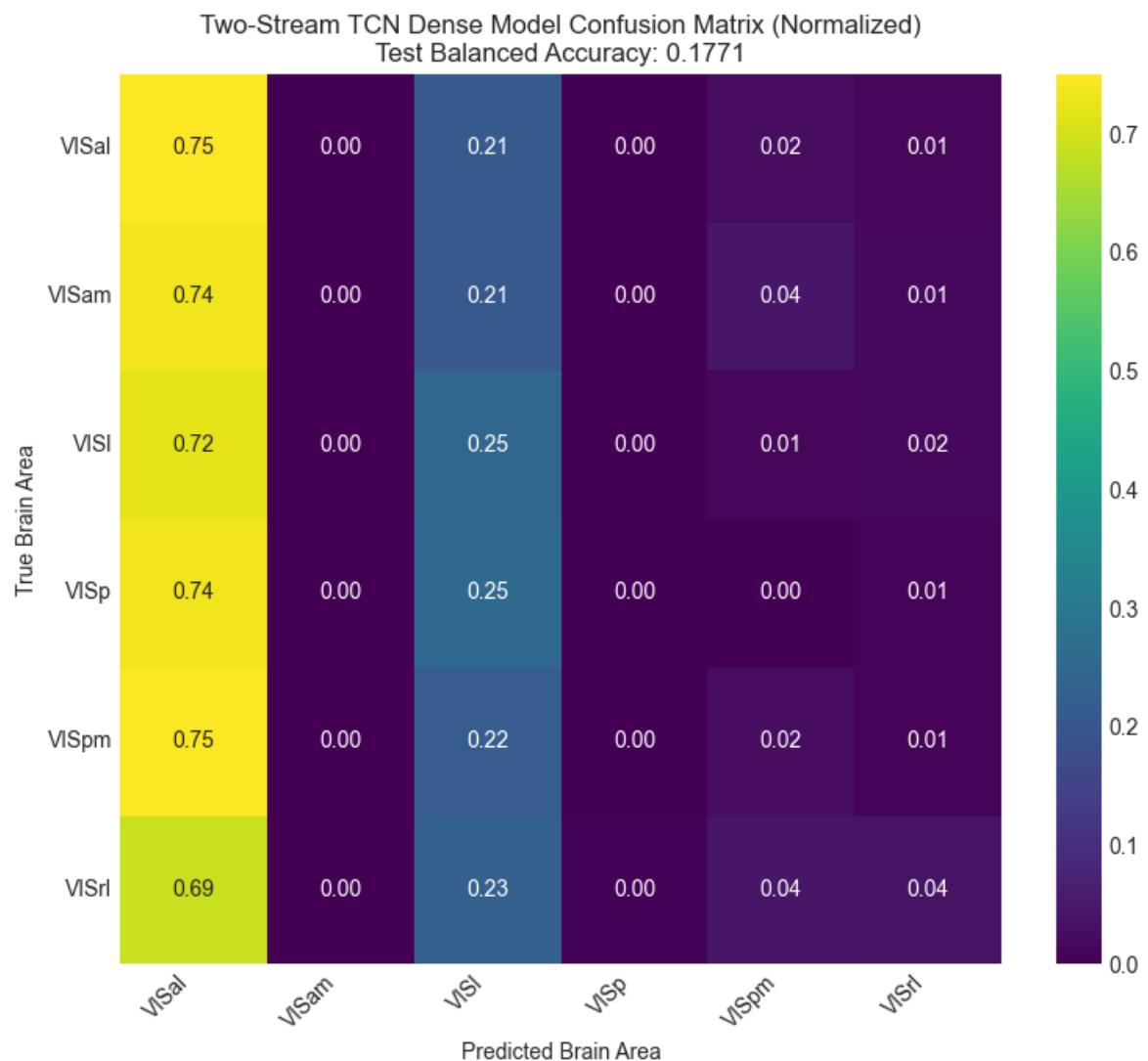
# LDA dense hippocampus



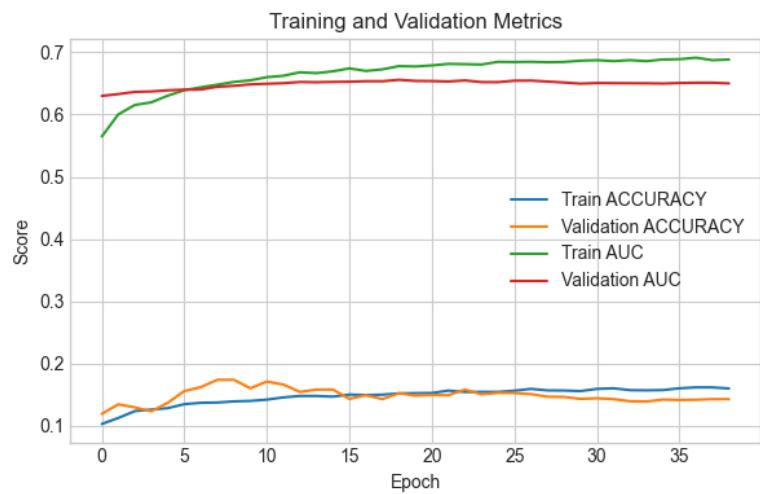
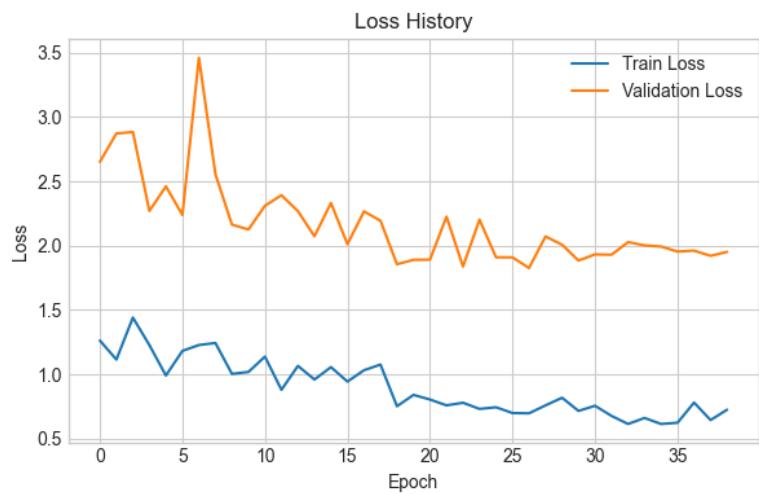
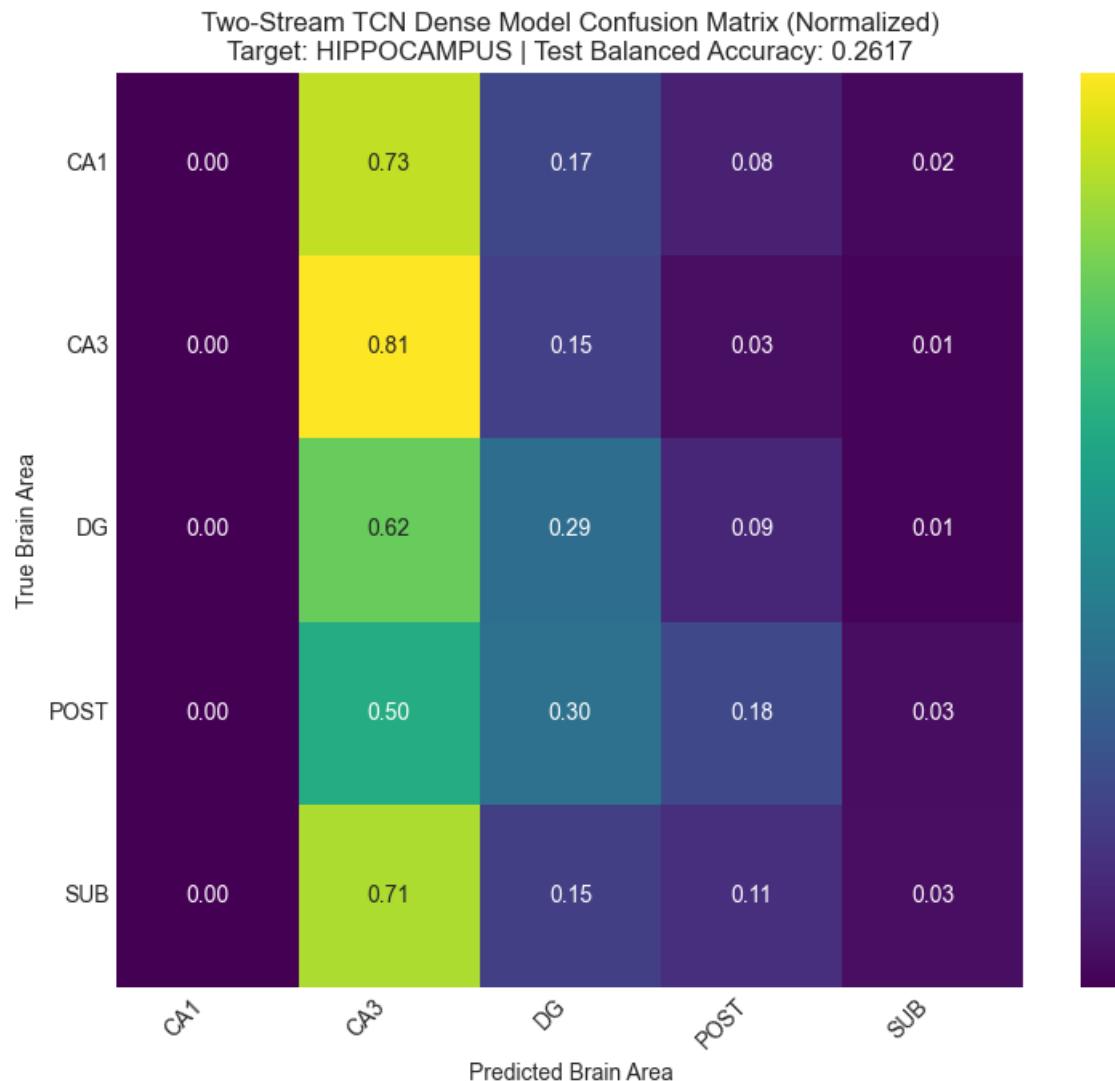
# LDA dense visual



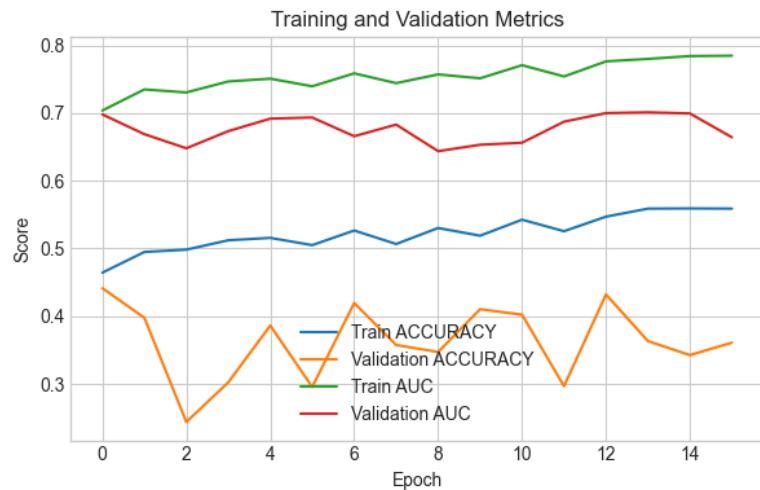
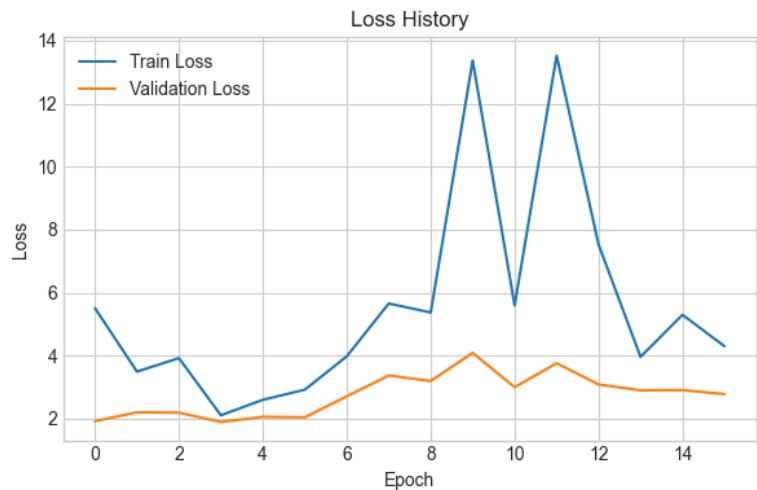
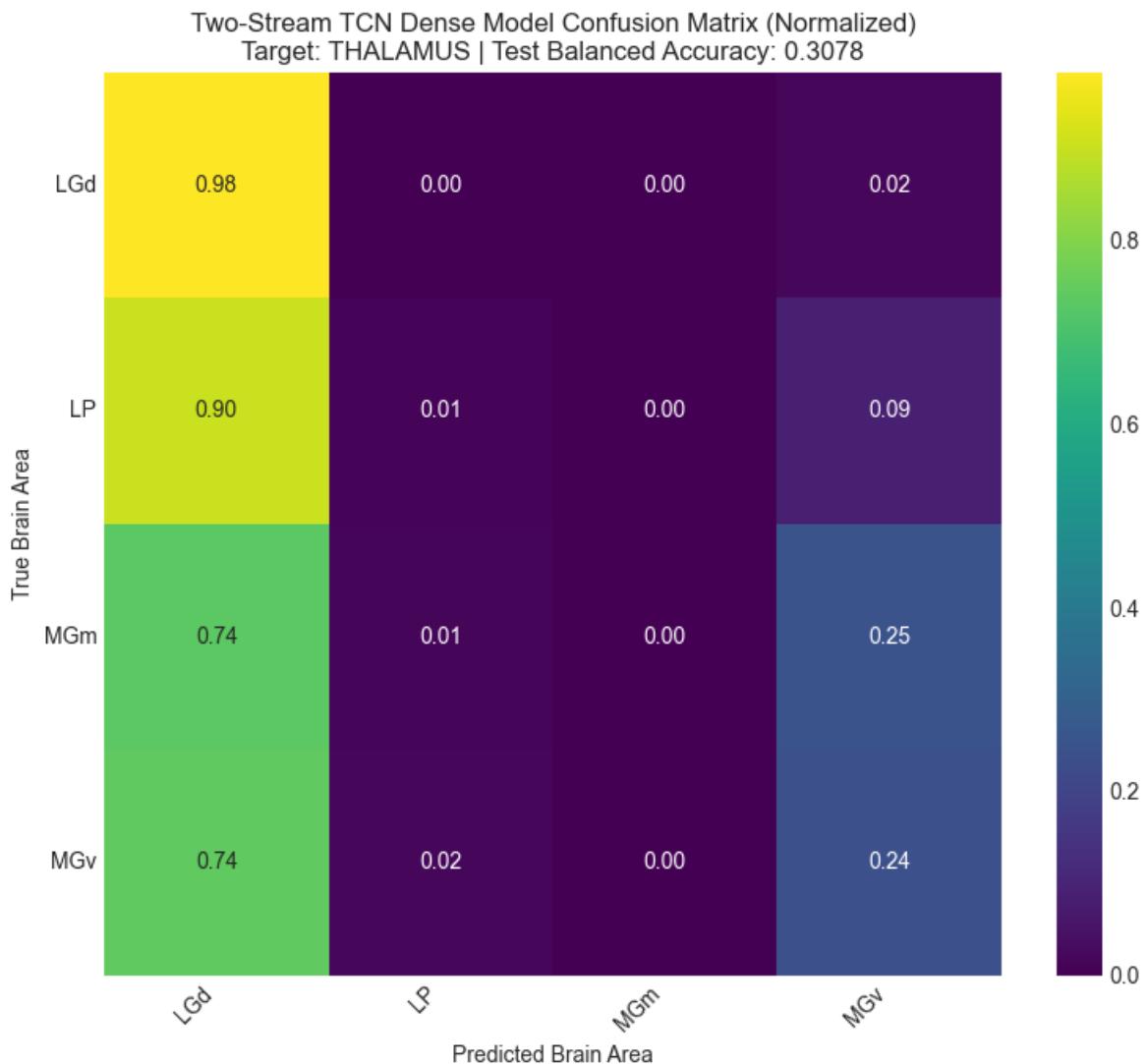
# TCN visual



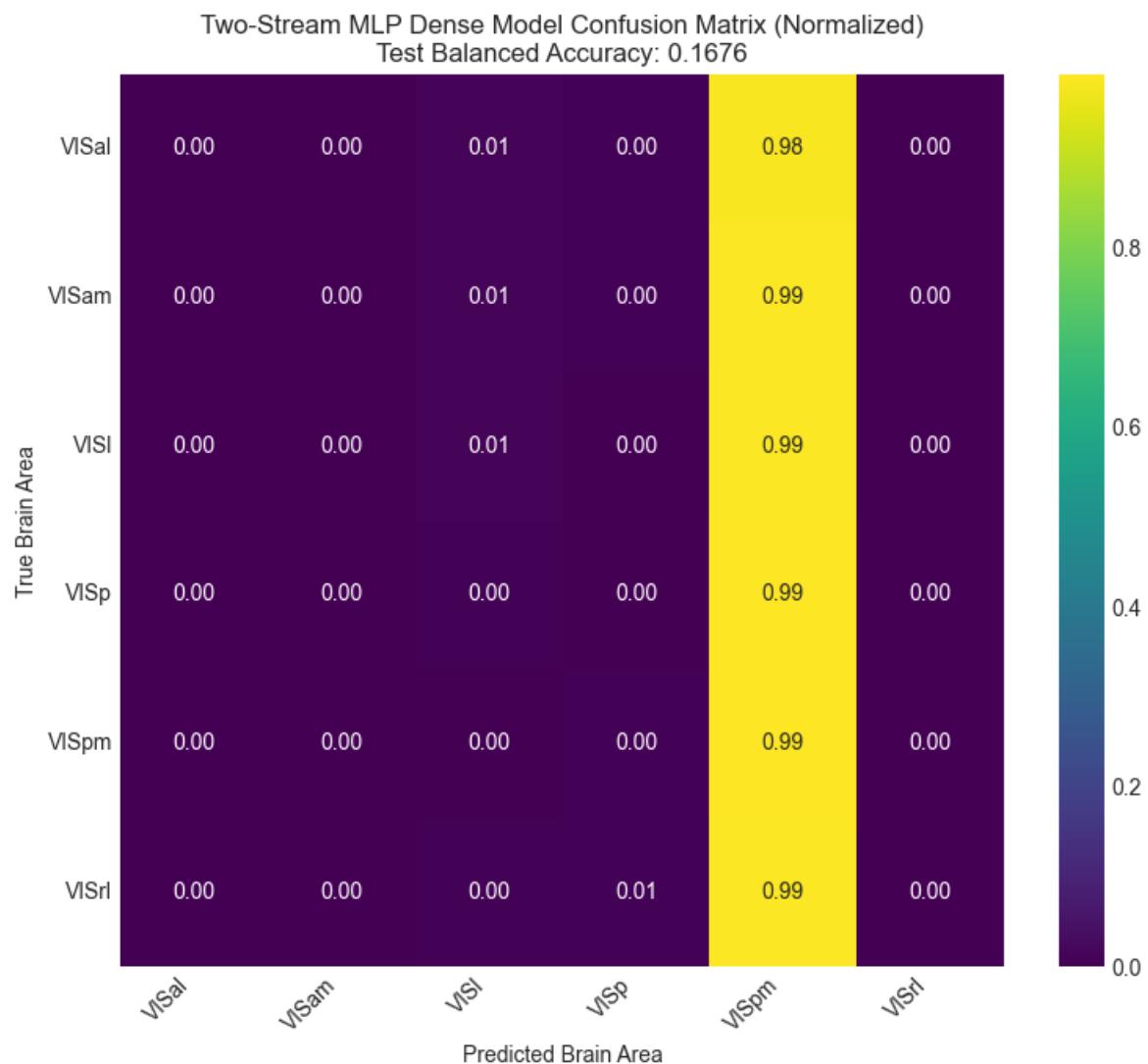
# TCN hippocampus



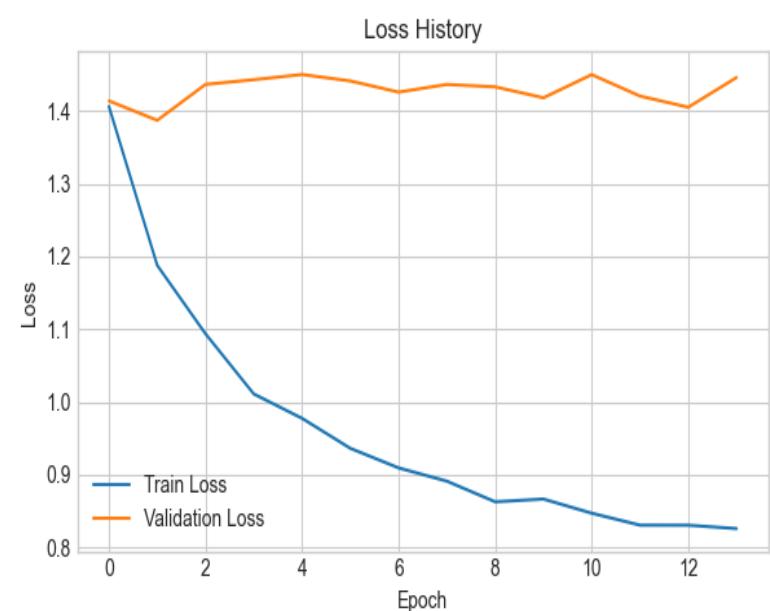
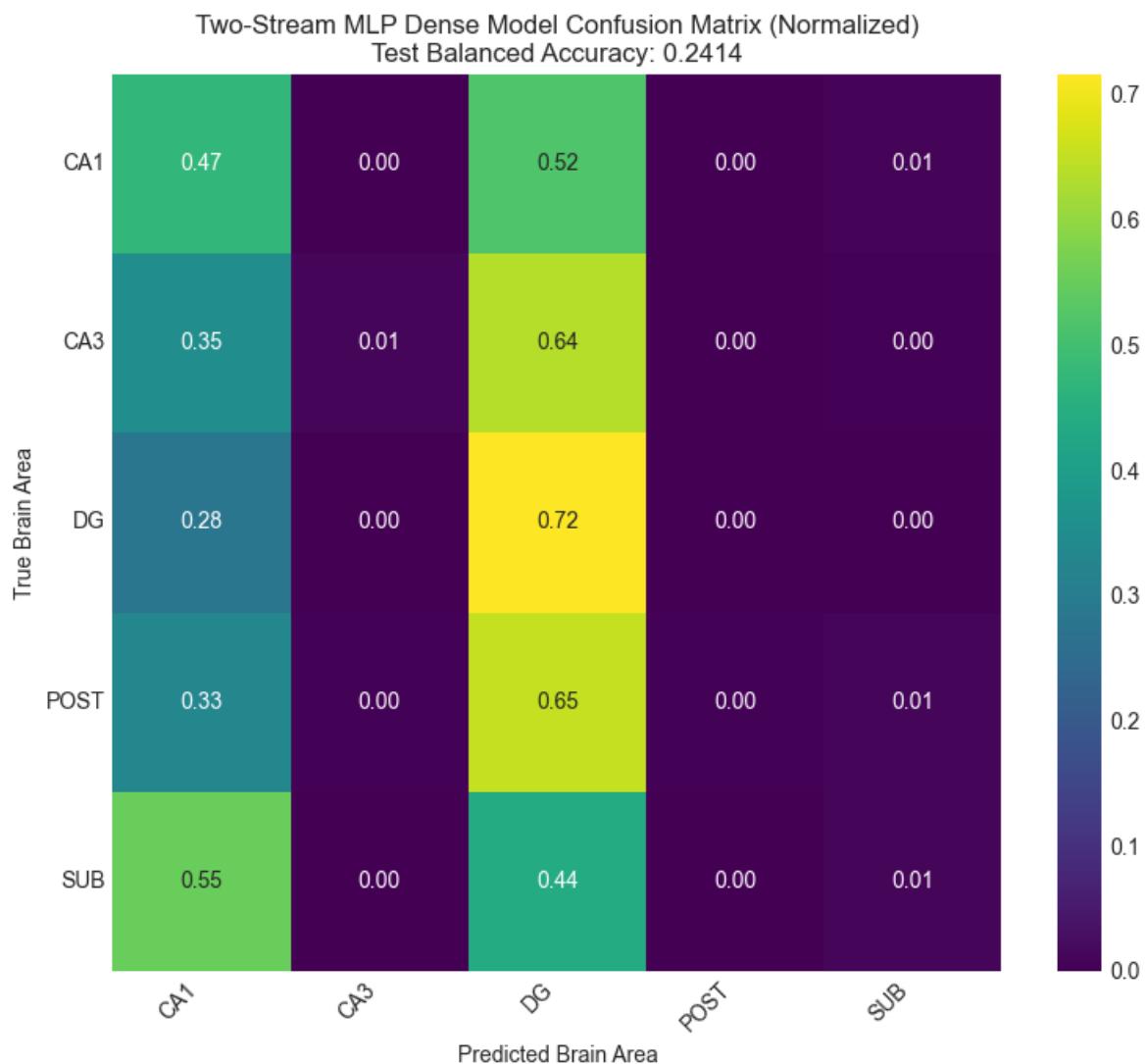
# TCN thalamus



# Dense MLP visual

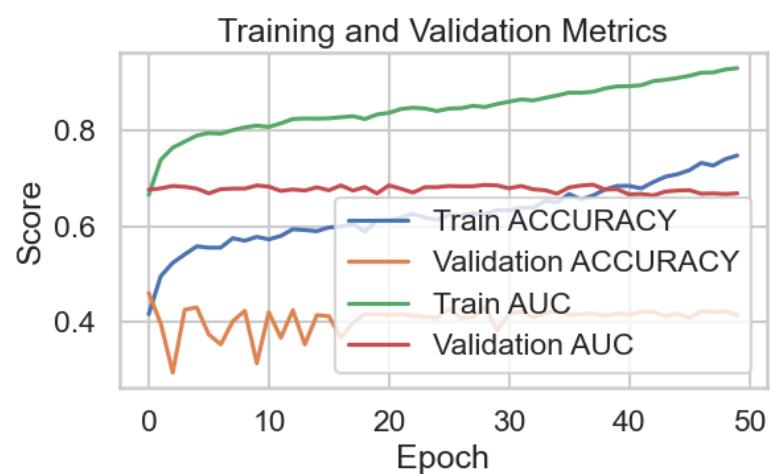
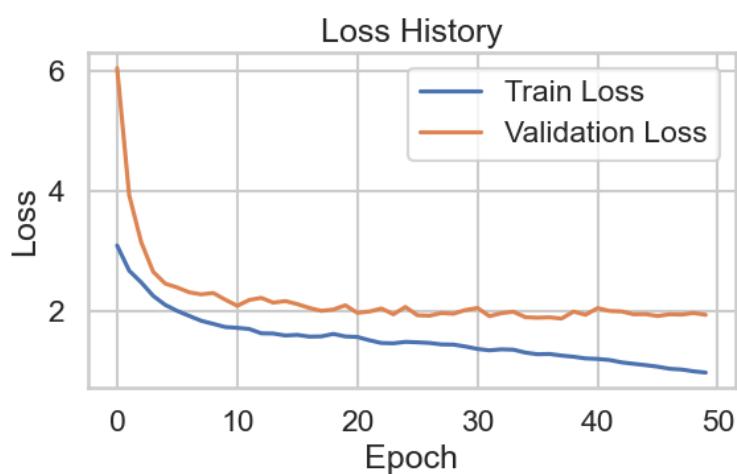
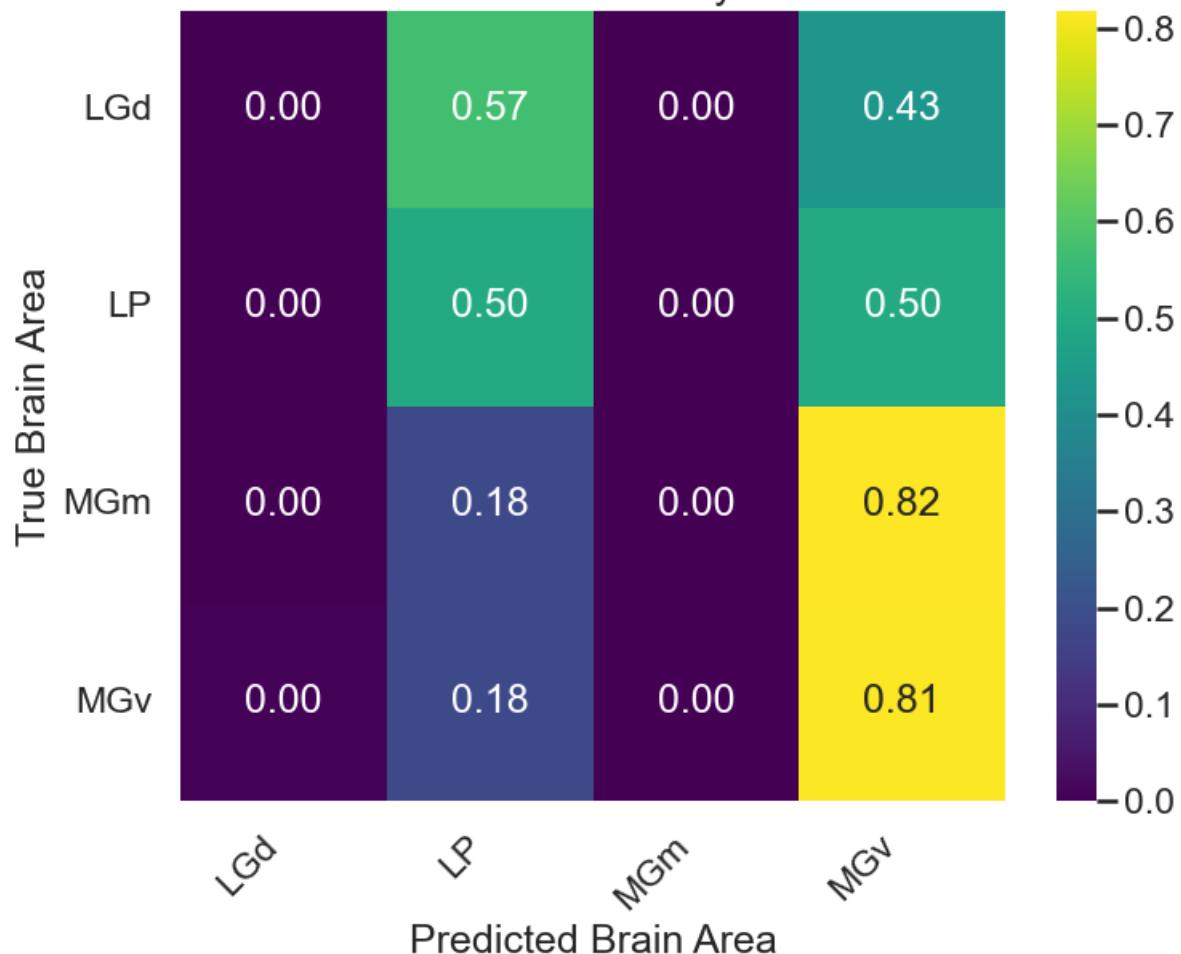


# Dense MLP hippocampus

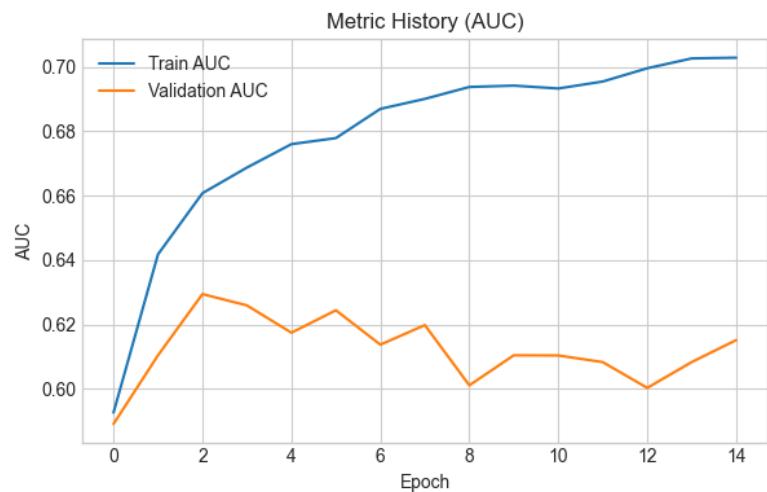
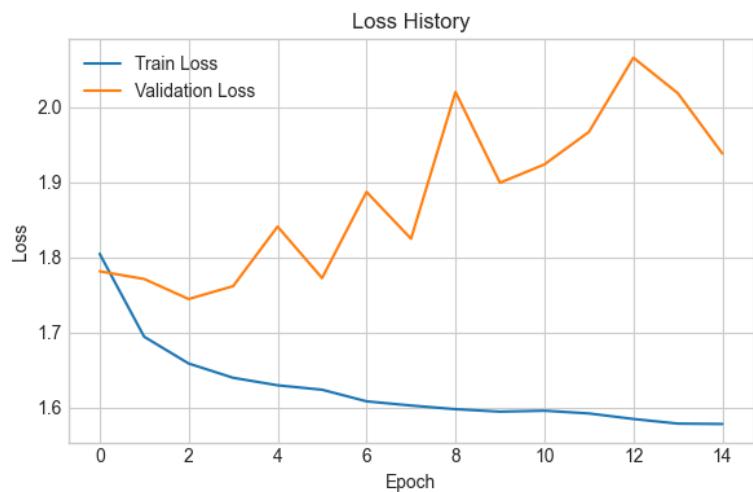
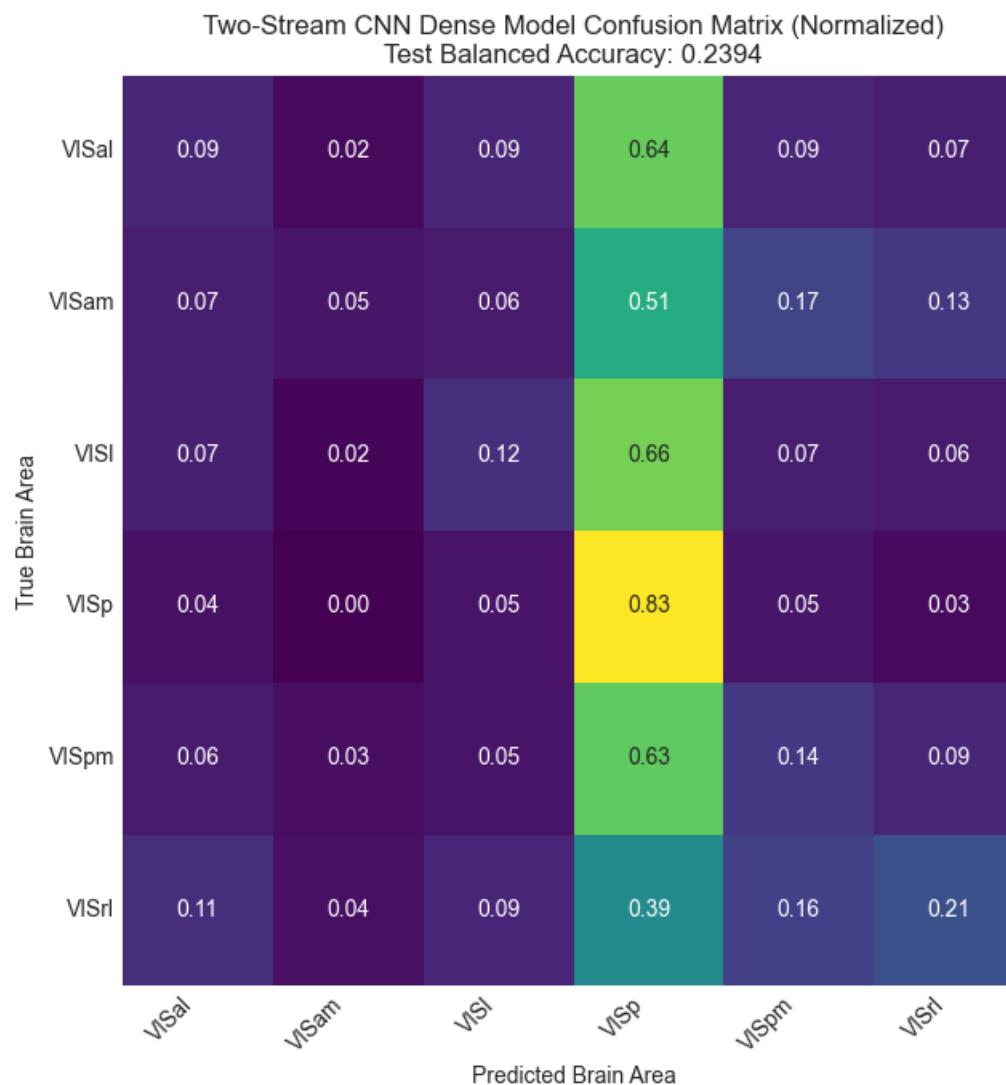


# Dense mlp thalamus

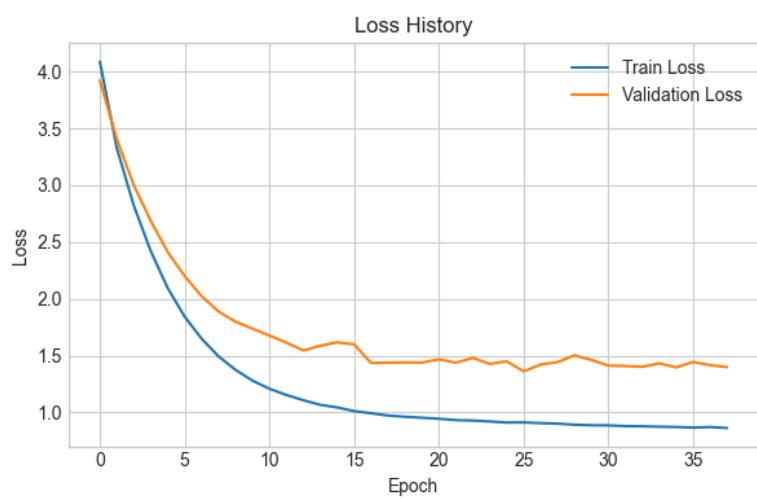
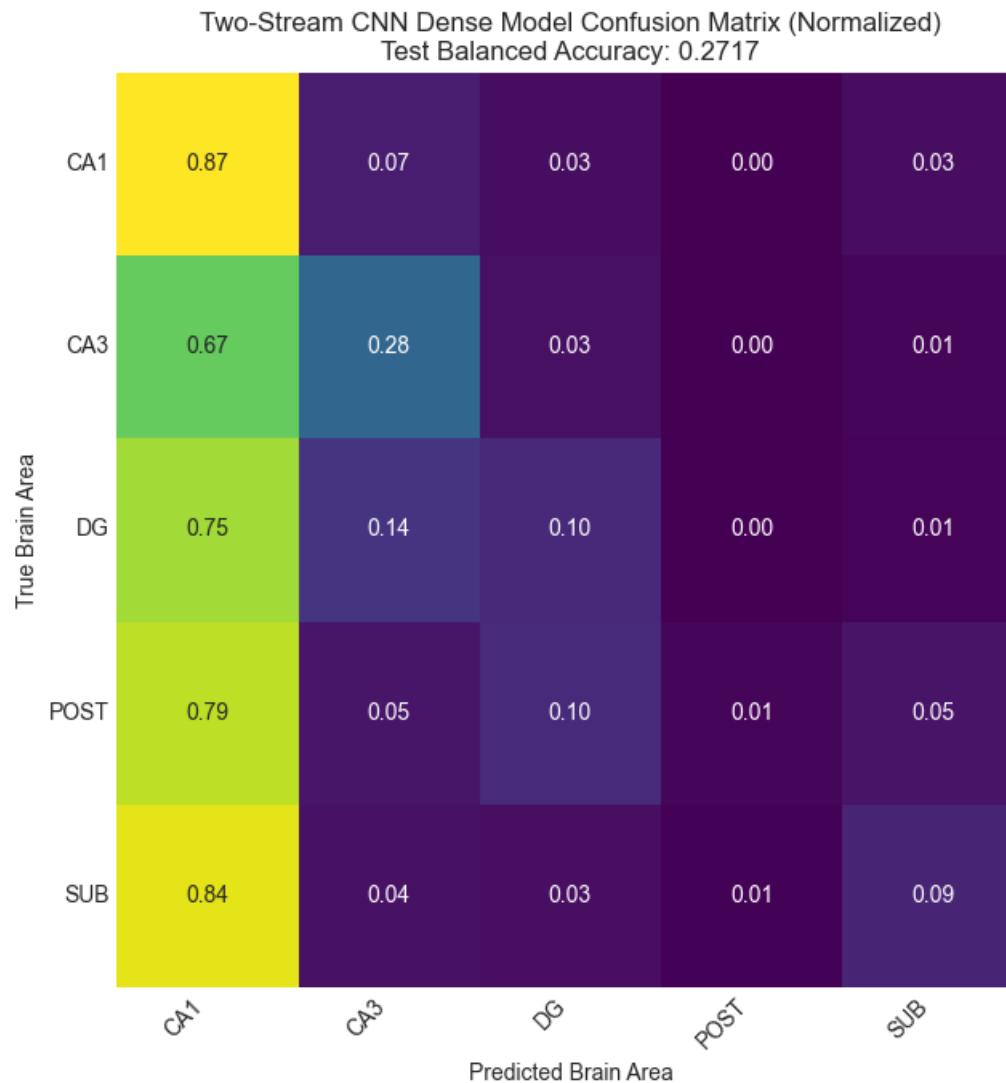
Two-Stream MLP Dense Model Confusion Matrix (Normalized)  
Test Balanced Accuracy: 0.3276



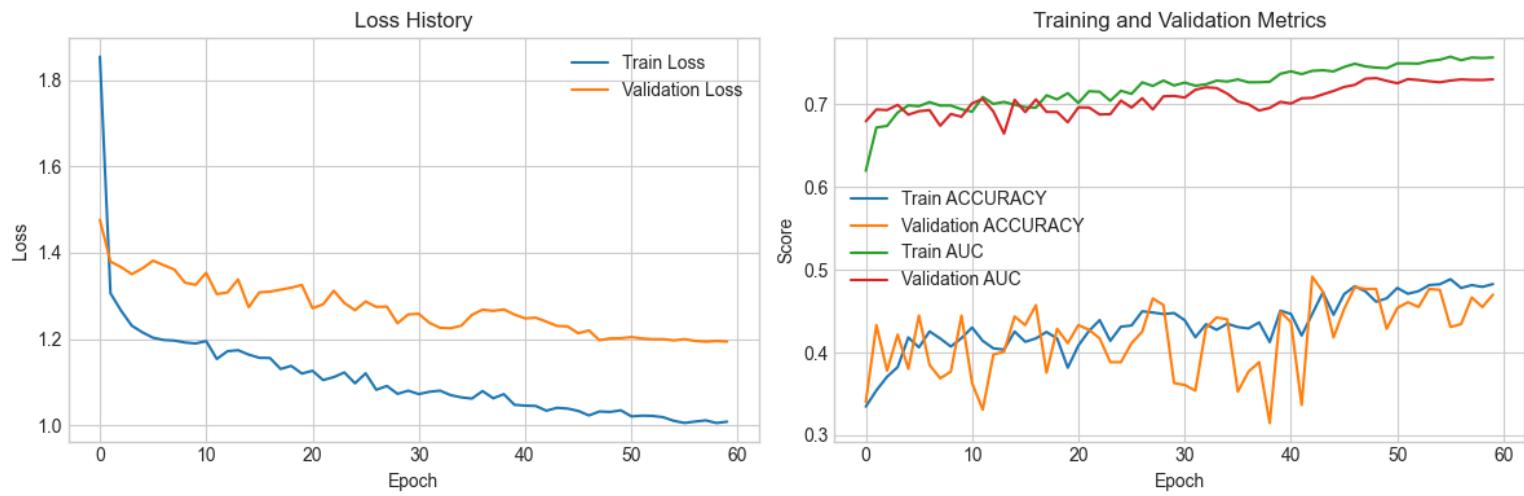
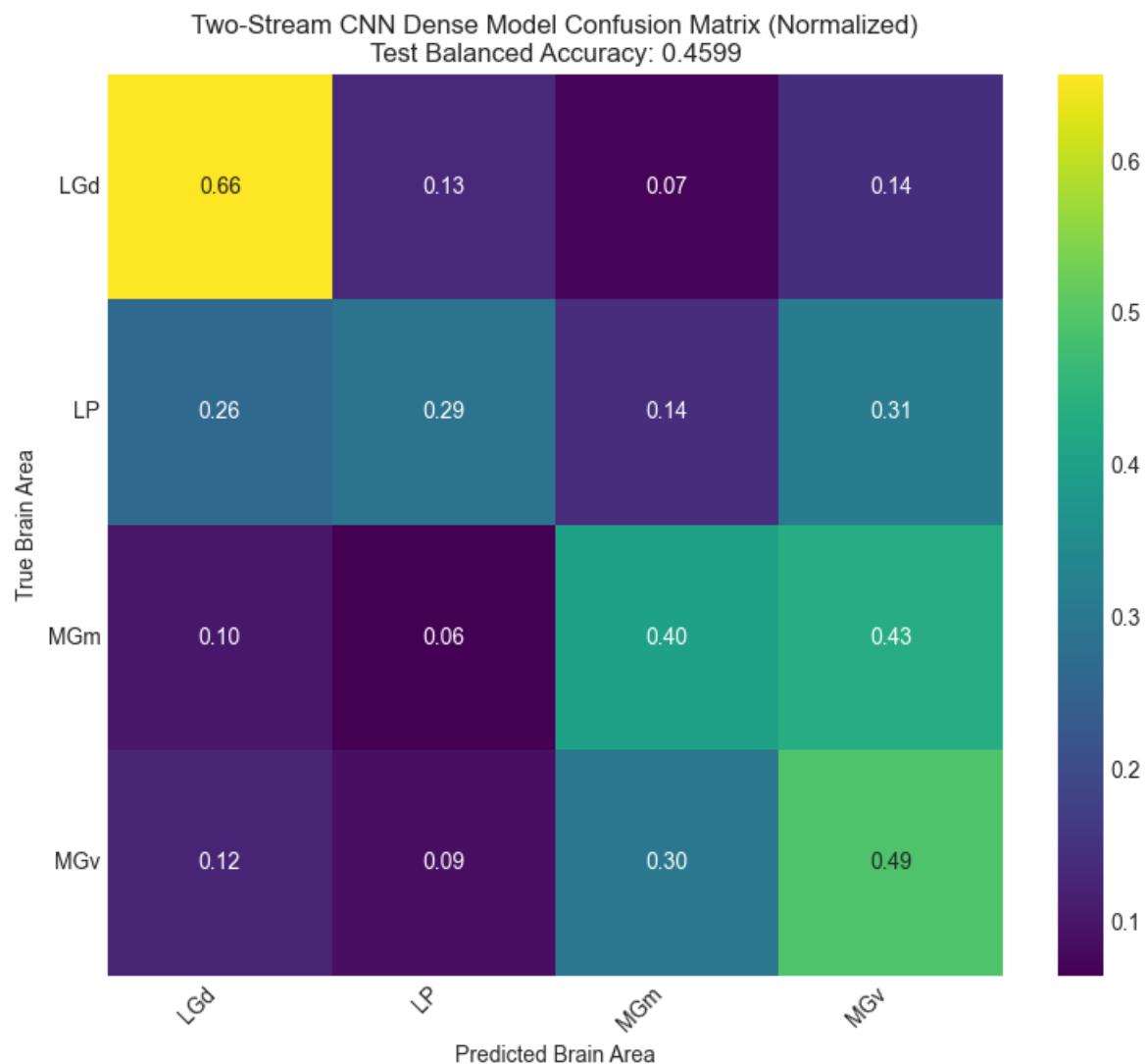
# CNN visual



# CNN hippocampus



# CNN thalamus



# Optimized hyper parameters(HP)

<b>Algorithm</b>	<b>Hyperparameter</b>	<b>All areas</b>	<b>Visual Areas Only</b>	<b>Hippocampus Only</b>	<b>Thalamus Only</b>
<b>DMLP</b>	Layer 1 units	1024	4096	1024	2048
	Layer 2 units	512	512	256	1024
	Scalar	80	96	96	64
	Head units	768	256	768	1024
	L2 regularization	$1.18 \times 10^{-3}$	$2.35 \times 10^{-3}$	$7.45 \times 10^{-6}$	$3.25 \times 10^{-4}$
	Starting dropout rate	0.25	0.4	0.65	0.5
	Learning rate	$1.42 \times 10^{-4}$	$2.78 \times 10^{-2}$	$4.91 \times 10^{-4}$	$4.50 \times 10^{-3}$
	Feature weight	4.1	6.6	0.9	1.5
	Weight dampening $\alpha$	1.5	0.4	0.9	0.4
<b>LDA Dense</b>	LDA layer units	128	128	448	256
	Scalar units	8	136	96	248
	Head units	1280	768	1024	768
	L2 regularization	$1.92 \times 10^{-5}$	$1.60 \times 10^{-6}$	$1.58 \times 10^{-3}$	$1.58 \times 10^{-3}$
	Starting dropout rate	0.4	0.30	0.15	0.15
	Learning rate	$2.21 \times 10^{-3}$	$1.15 \times 10^{-3}$	$1.77 \times 10^{-3}$	$5.61 \times 10^{-5}$
	Feature weight	7.4	8.7	6.5	7.9
	Weight dampening $\alpha$	0.45	1.45	0.60	1.00
	CNN layers	2	2	3	1
<b>CNN</b>	Filters	128	96	96	64
	Kernel size	10	5	20	20
	Pool size	2	4	4	4
	Scalar units	16	48	112	80
	Head units	256	1024	512	512
	L2 regularization	$6.02 \times 10^{-5}$	$1 \times 10^{-6}$	$6.51 \times 10^{-3}$	$1.5 \times 10^{-4}$
	Starting dropout rate	0.45	0.25	0.4	0.4
	Learning rate	0.482	$4.42 \times 10^{-4}$	$2.26 \times 10^{-4}$	$2.40 \times 10^{-2}$
	Feature weight	9.2	6.9	8.3	8.1
	Weight dampening $\alpha$	0.90	0.4	0.7	—
<b>Transformer</b>	Embedding dim	32	96	64	64
	Num. heads	6	6	8	2
	FF dim	512	128	256	384
	Num. blocks	1	2	1	1
	Scalar units	96	32	16	128
	Head units	1024	1024	768	512
	L2 regularization	$3.46 \times 10^{-4}$	$7.80 \times 10^{-4}$	$4.88 \times 10^{-5}$	$5.52 \times 10^{-5}$
	Starting dropout rate	0.10	0.15	0.10	0.10

<b>Algorithm</b>	<b>Hyperparameter</b>	<b>All areas</b>	<b>Visual Areas Only</b>	<b>Hippocampus Only</b>	<b>Thalamus Only</b>
<b>TCN</b>	Learning rate	$1.02 \times 10^{-3}$	$2.40 \times 10^{-4}$	$2.01 \times 10^{-2}$	$1.58 \times 10^{-3}$
	Feature weight	8.9	5.4	4.6	1.1
	Weight dampening $\alpha$	0.60	1.60	1.20	1.10
	Kernel size	5	5	5	5
	Num. filters	64	16	16	64
	Num. levels	8	6	5	4
	Scalar units	16	80	32	96
	Head units	1024	768	256	1024
	L2 regularization	$1.08 \times 10^{-6}$	$4.51 \times 10^{-6}$	$2.94 \times 10^{-6}$	$4.25 \times 10^{-3}$
	Starting dropout rate	0.1	0.55	0.30	0.35
<b>TCN</b>	Learning rate	$2.16 \times 10^{-3}$	$1.87 \times 10^{-4}$	$1.94 \times 10^{-4}$	$3.71 \times 10^{-3}$
	Feature weight	5.2	0.8	1.9	1.9
	Weight dampening $\alpha$	1.1	0.3	1.6	0.3