

Article

Not peer-reviewed version

---

# In Vitro to In Vivo: Bidirectional and High-Precision Generation of In Vitro and In Vivo Neuronal Spike Data

---

[Masanori Shimono](#)\*

Posted Date: 11 March 2026

doi: 10.20944/preprints202603.0875.v1

Keywords: in vitro; in vivo; neuronal spike trains; Transformer; Dice loss; focal loss; domain transfer; translational neuroscience



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# In Vitro to In Vivo: Bidirectional and High-Precision Generation of In Vitro and In Vivo Neuronal Spike Data

Masanori Shimono

School of Medicine, Keio University, Tokyo, Japan; smn@keio.jp

## Abstract

Translational neuroscience relies on both in vitro slice recordings and in vivo recordings. Their spontaneous population dynamics are observed under decisively different conditions, and across independent experiments there is typically no clear neuron-to-neuron correspondence. Here we formulate a time-resolved, bidirectional transfer task between in vitro and in vivo multineuronal spike trains and provide a standardized evaluation procedure for generation across markedly different recording preparations. We train an autoregressive Transformer on 1-ms binned, 128-unit binary sequences and introduce Dice loss to directly optimize spike-event overlap under extreme class imbalance, comparing it with Binary Focal Cross-Entropy ( $\gamma = 2.0$ ). Across 12 mouse datasets (6 in vitro HD-MEA sessions and 6 in vivo Neuropixels sessions), the method achieves strong within-domain performance and remains above chance for cross-domain generation (ROC-AUC  $0.70 \pm 0.09$  for in vitro  $\rightarrow$  in vivo;  $0.80 \pm 0.10$  for in vivo  $\rightarrow$  in vitro). Because spike events are rare, we report Precision–Recall curves and PR-AUC alongside ROC-AUC to reflect minority-event quality. To our knowledge, this is the first demonstration of bidirectional, time-resolved generation between unpaired in vitro and in vivo population spike trains without assuming cell correspondence, and the framework can be adapted to other sparse neural event data and related event-based datasets when domain-specific validation criteria are defined.

**Keywords:** in vitro; in vivo; neuronal spike trains; Transformer; Dice loss; focal loss; domain transfer; translational neuroscience

## 1. Introduction

The nervous system transmits signals and performs computation through the spiking activity of neurons. To understand how information is encoded at the level of large populations, it is essential to characterize the spatiotemporal “recipes” that emerge from interactions among many neurons.

Even without explicit cognitive tasks, the brain exhibits spontaneous activity. Across development, accumulated experience and internal constraints shape this ongoing activity, embedding a repertoire of latent patterns that can be rapidly recruited when external inputs arrive [1,2]. Yet precisely because spontaneous activity is rich and not time-locked to an external event, it has often been treated as “noise,” and its moment-to-moment evolution remains challenging to predict and compare across settings [1,3]. In this sense, spontaneous activity may carry a compressed trace of an individual’s past experience, while remaining difficult to model because of its complexity.

Population dynamics are not arbitrary: they are constrained by anatomical wiring and mesoscale organization. Large-scale anatomical and functional connectomics has sharpened the link—close but nontrivial—between structure and activity in cortex [4–7]. Complementary network-level analyses further quantify how structure–function coupling varies across space and time, reinforcing the view that spontaneous dynamics are shaped by an anatomical scaffold [6,7]. Moreover, effective interactions can sometimes be inferred from activity time series, implying that partially predictable dynamical motifs may be embedded in spontaneous fluctuations [8,9].

From a translational viewpoint, a central challenge is whether dynamical regularities learned under one recording preparation generalize to another. Compared with within-preparation settings emphasized in prior neural sequence modeling, transfer between in vitro and in vivo confronts a larger, compound shift: brain state and nonstationarity differ, noise and artifact structure (including spike-sorting biases) change, firing rates and sparsity regimes shift, and there is typically no neuron-to-neuron correspondence or paired trials to anchor the mapping. Nonetheless, in vitro preparations allow controlled manipulation and stable recording conditions, whereas in vivo recordings capture dynamics in an intact, behaving organism; quantifying links between these regimes could provide a reusable bridge for downstream analyses and inform coordinated experimental design [10,11].

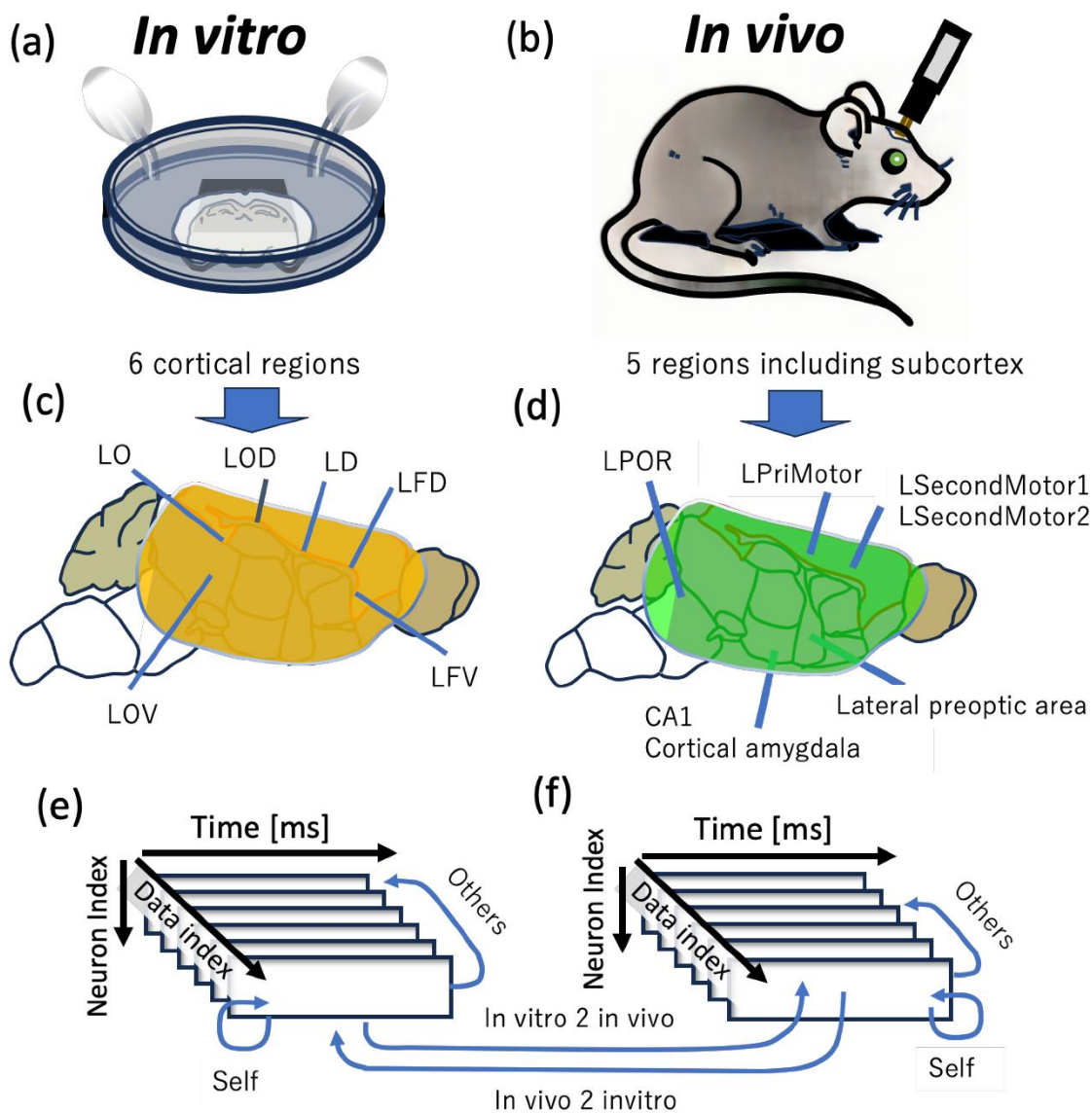
Recent sequence models now capture population spiking at scale, suggesting that transferable dynamical motifs exist even without explicit stimuli [12]. Transformer-based neural-data models (e.g., NDT, STNDT, NDT2) further enable efficient parallel modeling of binned spike trains, and are often evaluated via reconstruction, forecasting, or decoding within a shared recording context (same preparation, subject/session, or closely related tasks) [13–16].

Cross-preparation transfer, however, remains comparatively under-quantified in a time-resolved setting, despite its potential to generalize beyond a single dataset or laboratory setting.

We summarize our contributions as:

- We formulate and benchmark a bidirectional, time-resolved transfer task between in vitro and in vivo population spike trains, and we describe a standardized evaluation procedure for cross-preparation generation.
- We present one concrete implementation using an autoregressive Transformer for sparse binary events and compare Dice loss with Binary Focal Cross-Entropy ( $\gamma = 2.0$ ) under extreme class imbalance.
- We report both ROC- and PR-based metrics (ROC-AUC, Precision–Recall curves, and PR-AUC) for prediction/generation under extreme sparsity. Detailed hyperparameter sweeps and compute–accuracy trade-offs are reported in the Supplementary Materials.
- We clarify data splitting/independence and the 128-neuron standardization procedure, and we provide supplementary materials for reproducibility and additional analyses.

We formulate and benchmark a bidirectional, time-resolved transfer task between in vitro and in vivo population spike trains, and we provide a standardized evaluation procedure that can be reused across preparations. As one concrete instantiation, we implement an autoregressive Transformer for sparse binary events and evaluate Dice loss against Binary Focal Cross-Entropy ( $\gamma = 2.0$ ). To quantify performance under severe sparsity, we report both ROC- and PR-based metrics (ROC-AUC, Precision–Recall curves, and PR-AUC) for prediction and generation. Finally, we clarify data splitting and independence assumptions, detail the 128-neuron standardization procedure, and provide supplementary materials to support reproducibility and additional analyses.



**Figure 1.** Illustration of the Data Sets Used in This Study.

Figure 1 provides an overview of the data sets utilized in this study. Panel (a) depicts the setup of an *in vitro* electrophysiological experiment. A brain acute slice is placed on an electrode, perfused with artificial cerebrospinal fluid (ACSF), and continuously bubbled with oxygen-enriched gas while neural activity is recorded. Panel (b) illustrates the setup for an *in vivo* electrophysiological experiment. In this case, electrodes are inserted into targeted brain regions of a living mouse to measure neural activity.

Panels (c) and (d) indicate the brain regions recorded during the *in vitro* and *in vivo* experiments, respectively. These panels display abbreviated names representing each recorded brain region. The meanings of these abbreviations and the specific brain regions they correspond to are summarized in Table 1.

Panels (e) and (f) represent two-dimensional neural activity data obtained from each brain region over time, with individual squares corresponding to neuronal activity at different time points. The blue arrows illustrate the generated pattern pairs within each experiment. Within both (e) and (f), the blue arrows indicate self-generation from the source data, as well as generation towards other regions (self-to-self and self-to-others). Additionally, inter-experimental generation pathways between *in vitro* and *in vivo* data are depicted, specifically *in vitro*-to-*in vivo* and *in vivo*-to-*in vitro* generation, as shown by the interconnecting blue arrows between panels (e) and (f).

The recorded brain regions and abbreviations used in this study are summarized in Table 1.

**Table 1.** List of Brain Regions Corresponding to Experimental Data.

|                     |                |                                  |   |
|---------------------|----------------|----------------------------------|---|
| <b>LO</b>           | shimonolab     | Left Occipital<br>210126         | Gustatory areas, cingulate area, Agranular insular area, Perirhinal area                    |
| <b>LOV</b>          | shimonolab     | Left Occipital<br>Ventral 210706 | Temporal association area,<br>Supplemental somatosensory area                               |
| <b>LFV</b>          | shimonolab     | Left Frontal Ventral<br>200615   | Primary motor area, Gustatory area, Orbital area,<br>Agranular insular area                 |
| <b>LFD</b>          | shimonolab     | Left Frontal Dorsal<br>190806    | Primary motor area, Secondary motor area,<br>Anterior cingulate area                        |
| <b>LD</b>           | shimonolab     | Left Dorsal<br>200609            | Primary somatosensory area, Anterior area,<br>Supplemental somatosensory area               |
| <b>LOD</b>          | shimonolab     | Left Occipital<br>Dorsal 190910  | Primary visual area, Dorsal auditory area,<br>Anterolateral visual area, Retrosplenial area |
| <b>LPOR</b>         | angelakilab    | NYU-40/2021-<br>04-14            | The postrhinal area, VISpor   |
| <b>LPriMotor1</b>   | mrsicflogellab | SWC_038/2020-<br>08-01           | Primary motor area  |
| <b>LSecMotor2</b>   | mrsicflogellab | SWC_038/2020-<br>07-31           | Secondary motor area  |
| <b>LSecMotor</b>    | mrsicflogellab | SWC_038/2020-<br>07-30/          | Secondary motor area  |
| <b>LLatPreopt</b>   | hoferlab       | SWC_043/2020-<br>09-21           | Lateral preoptic area   |
| <b>LHippCorAmyg</b> | hoferlab       | SWC_043/2020-<br>09-20           | Hippocampus, field CA1 + Cortical amygdala<br>area posterior part lateral zone (COAP)       |

Table 1 presents a comprehensive list of brain regions corresponding to the in vitro and in vivo experimental data. From left to right, the table includes: (1) the dataset name as used in this study, (2) the laboratory responsible for data acquisition, (3) supplementary dataset information such as measurement dates, and (4) the specific brain regions recorded.

The in vitro data were obtained by the Shimono Lab, with recordings conducted from multiple regions of the left cerebral hemisphere. In contrast, the in vivo data were collected by the AngelakiLab, Mrsic-FlogelLab, and Hofer Lab, covering a wide range of brain regions, including the motor cortex, visual cortex, hippocampus, and amygdala.

For further reference, see: [https://viz.internationalbrainlab.org/app?spikesorting=ss\\_2024-05-06&dset=bwm&pid=9915765e-ff15-4371-b1aa-c2ef1db8a254&tid=5&cid=388&qc=1](https://viz.internationalbrainlab.org/app?spikesorting=ss_2024-05-06&dset=bwm&pid=9915765e-ff15-4371-b1aa-c2ef1db8a254&tid=5&cid=388&qc=1).

Notably, for both the in vitro and in vivo collections, each dataset was recorded from a different mouse, so no individual animal contributes data to more than one dataset, ensuring independence across datasets.

## 2. Materials and Methods

### 2.1. Data Acquisition

#### 2.1.1. In Vitro Data

The in vitro data were collected from cortical slices of C57BL/6 mice (3–5 weeks old) using high density microelectrode arrays (HD MEAs, MaxWell Biosystems Co.), following the methodologies described in Nakajima et al. [12] and Matsuda et al. [31]. All slices were cut orthogonal to the cortical surface, and the slice orientation documentation is summarized in the referenced Extended Data (Matsuda et al. [31], Extended Data Figure 1-1). For the MaxOne system, the distance between adjacent electrodes is 15  $\mu\text{m}$ , providing a fixed spatial sampling grid across sessions.

Prior to tissue extraction, each mouse was deeply anaesthetized with 1–1.5% isoflurane delivered via inhalation. While full surgical anesthesia was confirmed (no pedal reflex), euthanasia was performed by cervical dislocation. The brain was then rapidly removed and immersed in ice cold cutting solution for slice preparation.

From the recorded electrical signals, spike sorting was carefully conducted using SpykingCircus software, allowing for the extraction of neural activity data from approximately 1,000 adjacent cells. To ensure a standardized input dimensionality without assuming cell to cell correspondence across datasets, we used 128 units per session. Specifically, we excluded the first 100 units and then selected the next 128 units consecutively in index order, which provides a deterministic and reproducible sampling rule while reducing potential edge related effects. The corresponding data loading and downsampling code is provided in the Supplementary Materials.

In previous studies, cortical data were categorized into 16 groups (eight per hemisphere). However, considering that in vivo data were exclusively recorded from the left hemisphere, we refined our dataset selection accordingly. To maintain consistency and data quality, we limited our analysis to six groups of in vitro data from the left hemisphere (Table 1 and Figure 1).

To minimize the impact of non stationarity, we excluded the first 30 minutes of recordings from each in vitro session. The subsequent 10.0 minutes were then segmented into 5.0 minutes for training, 2.5 minutes for validation, and 2.5 minutes for testing. A detailed list of the in vitro datasets used in this study is provided in Table 1. For further details regarding the dataset, refer to the cited studies [12,31]. Where additional slice level metadata are available, we report them; otherwise, we clarify the limits of quantifying slice to slice differences within the current dataset release.

#### 2.1.2. In Vivo Data

The in vivo data were obtained from the International Brain Laboratory (IBL), specifically using passive stimulation protocol recordings from C57BL/6 mice (15–63 weeks old) [18]. The in vivo data used in this study were recorded between 111 to 442 days of age (mean: 34.43 weeks, median: 26.0 weeks), based on data collected before 2022. For analysis, we focused on 10.0 minutes of spontaneous activity recorded at the beginning of each session.

The International Brain Laboratory compiles data from multiple laboratories worldwide. In this study, we selected datasets containing spontaneous activity recorded in several of these laboratories for analysis. The full list of datasets used is summarized in Table 1.

All electrophysiological recordings in this dataset were collected using Neuropixels1.0 multi-electrode probes [19]. Spike sorting was performed using a motion-corrected three-dimensional spike localization method optimized for this electrode type [20]. Further details on experimental procedures and data processing pipelines for extracting neural activity time series can be found in the IBL experiment documentation:

[https://int-brain-lab.github.io/iblenv/notebooks\\_external/loading\\_passive\\_data.html](https://int-brain-lab.github.io/iblenv/notebooks_external/loading_passive_data.html)

## 2.2. Analysis Methods

Study scope and integration. Biologically, we analyze spontaneous population spiking recorded under distinct preparations (in vitro slices vs. in vivo Neuropixels sessions; Section 2.1). Algorithmically, we cast each session as a 1-ms binned 128-unit binary sequence and train an autoregressive Transformer to predict/generate future bins under within- and cross-domain distribution shifts (Sections 2.2.1–2.2.5). Our primary contribution is this predictive-generation framework and its controlled evaluation; we do not claim discovery of new biological mechanisms. Interpretability analyses (attention maps and attention-weighted importance) are reported as exploratory probes of model behavior rather than definitive biological explanations (Sections 2.2.3–2.2.4; Results 3.3).

### 2.2.1. Data Segmentation

In both the in vitro and in vivo datasets shown in Figure 2a and Figure 2b, we extracted 10.0 minute segments of spontaneous activity. Spike times were discretized into binary event sequences using 1 ms time bins, with  $\Delta t$  equal to 1 ms, allowing at most one spike event per bin per unit. This yields an input tensor with shape time bins by units. The original acquisition sampling frequency and the spike sorting pipeline are dataset specific and are summarized in Section 2.1, together with the corresponding references where available.

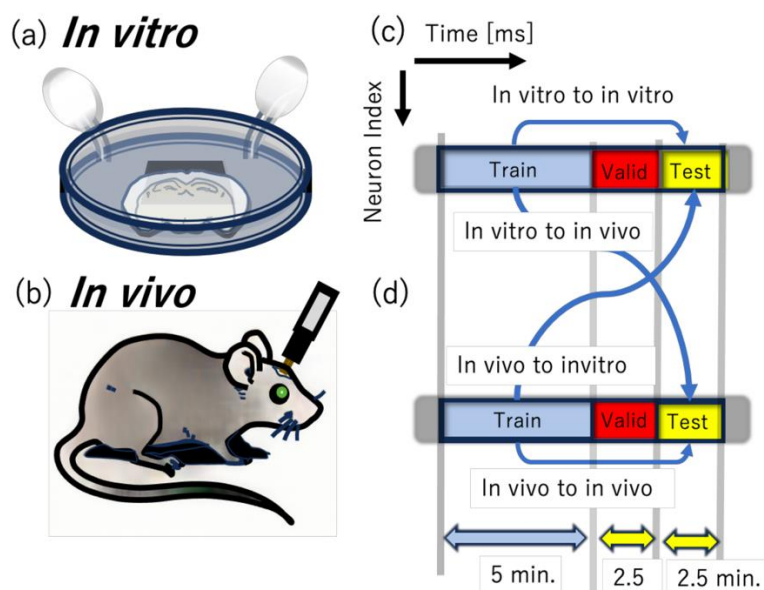
For all experiments, we standardized each session to 128 units to avoid assuming any one to one cell correspondence across datasets. While we confirmed that performance is not sensitive to random unit selection, we use a simple deterministic procedure to ensure reproducible and spatially balanced sampling in both in vitro and in vivo recordings, where electrode indices reflect spatial arrangement. Specifically, to reduce edge related effects and to obtain representative coverage across cortical layers, we discard the first 100 units and then select the next 128 units consecutively in index order. We describe this procedure in the preprocessing section and provide the corresponding data loading and downsampling code in the Supplementary Materials.

Each 10.0-minute recording was split into three non-overlapping segments: a 5.0-minute training set, a 2.5-minute validation set, and a 2.5-minute test set (Figure 2c). The training set was used to fit model parameters, the validation set was used for model selection and hyper-parameter tuning, and the test set was reserved for a single final evaluation. This split enables an assessment of generalization without overlap or double dipping.

For within-domain generation, such as in vitro to in vitro and in vivo to in vivo, training, validation, and test data were drawn from a single continuous 10.0-minute recording session and split into non-overlapping segments: 5.0 minutes for training, 2.5 minutes for validation, and 2.5 minutes for testing.

For cross-domain generation, such as in vitro to in vivo, training, validation, and test were taken from entirely different segments and, when applicable, different recordings or datasets, again using non-overlapping splits of 5.0 minutes for training, 2.5 minutes for validation, and 2.5 minutes for testing. Hyper-parameter selection and early stopping, when used, are performed only on the validation split, and the test split is used once for the final report, preventing overlap or double dipping in evaluation.

Our framework does not employ iterative re-training on generated outputs, not due to concerns about model collapse, but because faithful reproduction of original data characteristics is central to enabling direct, dynamics-preserving comparisons across datasets.



**Figure 2.** This figure explains the time duration of extracted data and the method of dividing data into training and test sets. (a) and (b) show *in vitro* data in the upper panels and *in vivo* data in the lower panels. For both (c) *in vitro* and (d) *in vivo* data, we extract 5-minute segments for training data and two 2.5-minute segments for validation and test data to conduct learning, validation and prediction. The combinations of choosing either *in vitro* or *in vivo* data for training and testing are classified as *in vitro* to *in vitro*, *in vitro* to *in vivo*, *in vivo* to *in vitro*, and *in vivo* to *in vivo*.

### 2.2.2. Transformer Model and Loss Function Selection

The primary analytical model employs a Transformer encoder [13] with multi-head self-attention (16 heads in the default configuration) to model binned population spike trains ( $\Delta t = 1$  ms) as sparse binary event sequences. We train autoregressively and use Dice loss [18] to directly optimize spike-event overlap under extreme class imbalance, and we compare against Binary Focal Cross-Entropy ( $\gamma = 2.0$ ). Hyperparameter selection (input length, depth, number of heads, and dropout) was guided by validation-set sensitivity analyses and a depth ablation reported in the Supplementary Materials (Text S1; Supplementary Figure S1 (validation) and Supplementary Figure S2 (generation)). In the *in vitro*→*in vitro* sensitivity analyses, input length, dropout, and the number of heads each exhibited a clear best-performing setting within the tested range, and we adopt the corresponding values as defaults (input length = 3,125 ms, dropout  $p = 0.1$ , heads = 16). For encoder depth, validation performance did not improve monotonically with depth, and deeper models increased computational cost; therefore, we adopt depth = 1 as the default for computational efficiency and stability, and we report the full sweeps for transparency.

### 2.2.3. Attention Map

To gain deeper insights into the information learned by the Transformer model during prediction generation, we quantitatively evaluated the model's learned information and predictive features by combining attention map dynamics with gradient-based importance. Here, we explain the relevant aspects of the Transformer model's internal structure.

The self-attention mechanism, which achieved breakthrough results particularly in Natural Language Processing (NLP) [13], is considered the central mechanism in Transformer model learning. At its core is the attention map, which represents a matrix indicating how much attention each input element (token) should pay to other input elements. The attention map is mathematically expressed as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \cdot V$$

Here  $Q$ ,  $K$ , and  $V$  denote the query, key, and value matrices computed from an input sequence  $X$  by linear projections:  $Q = X \cdot W_Q$ ,  $K = X \cdot W_K$ , and  $V = X \cdot W_V$ , where  $W_Q$ ,  $W_K$ , and  $W_V$  are learned weight matrices and  $(\cdot)$  denotes matrix multiplication.  $K^T$  denotes the transpose of  $K$  and  $d_k$  is the key dimension used for scaling.

In this work,  $X$  represents a sequence of binary spike vectors over time ( $\Delta t = 1$  ms) for 128 units.

The similarity scores  $S$  are scaled by  $\sqrt{d_k}$  and normalized with softmax to yield the attention weights  $A$ . These weights are then used to compute a weighted sum of  $V$  to produce the final output.

While attention maps have been considered fundamental to success in various tasks including machine translation [13], document summarization [21], and image recognition [14], recent studies have indicated that attention maps alone are insufficient to explain transformer learning [22,23]. Indeed, our research revealed phenomena that cannot be explained by attention maps alone, leading us to calculate the model's gradient-based importance.

#### 2.2.4. Gradient-Based Importance and Attention-weighted Importance

Gradient-based importance measures how strongly each input element influences the loss. Let  $L$  denote the training loss and  $h_{i,k}$  be the  $k$ -th feature of the model input (or embedding) at time-bin index  $i$ . We define the element-wise importance as  $g_{i,k} = |\partial L / \partial h_{i,k}|$  and the token-level importance as  $G_i = \sum_k g_{i,k}$ .

In other words,  $G_i$  aggregates gradient magnitudes over feature dimensions and indicates which input time bins/units most affect the loss. This provides a global importance score that is complementary to attention distributions.

To relate importance to pairwise interactions captured by attention, we combine  $G$  with the attention matrix  $A$  (query-to-key weights) to compute an attention-weighted importance:  $AWI_{i,j} = A_{i,j} \cdot G_j$ . This quantifies how much query position  $i$  attends to a key position  $j$  that is also influential for the loss.

We use AWI to visualize informative interactions during prediction and generation, and we report summary statistics in the Results.

Note that we do not place these terms in quotation marks; throughout the manuscript we use 'attention matrix' and 'importance' in the standard sense.

#### 2.2.5. Generation Data Evaluation Method

During testing, we generated one-step-ahead predictions at each 1-ms time bin ( $\Delta t = 1$  ms) for all 128 units, using the observed past test sequence as input while keeping the trained model fixed. We did not feed predicted spikes back into the input stream during evaluation. To evaluate the predicted spike trains, we report ROC-AUC and Precision-Recall metrics including PR-AUC and average precision, together with firing-rate and inter-spike-interval statistics. Under extreme sparsity, PR curves and PR-AUC are emphasized for spike-event quality because ROC-AUC can be overly optimistic when negatives dominate.

#### 2.2.6. Network Layout Optimization in Three-Dimensional Space

We mapped recording sites and regions in three-dimensional space so that Euclidean distance reflects transfer difficulty between datasets. For each ordered pair of nodes  $i$  and  $j$ , we computed a directed transfer score  $score_{i \rightarrow j}$  (e.g., ROC-AUC of predicting dataset  $j$  from dataset  $i$ ). We then formed a symmetric similarity  $s_{ij} = 0.5 \cdot (score_{i \rightarrow j} + score_{j \rightarrow i})$  and converted it to a dissimilarity by the inverse transform  $d_{ij} = s_{ij}^{-1} - 1$ , so that perfect transfer maps to  $d_{ij} = 0$  and chance-level performance (ROC-AUC = 0.5) maps to  $d_{ij} = 1$ . Given the target dissimilarities  $\{d_{ij}\}$ ,

we optimized node positions  $x_i \in \mathbb{R}^3$  by minimizing the stress energy  $E(x) = \sum_{\{i < j\}} (||x_i - x_j|| - d_{ij})^2$ , initialized from random positions and optimized by gradient descent until convergence (Figure 4f). This layout is intended as an exploratory visualization and should not be interpreted as an anatomical distance metric.

### 2.3. Ethics Statement

Ethical approvals and compliance statements for the in vitro and in vivo experiments will be provided in the Institutional Review Board Statement (Back matter), following MDPI requirements.

## 3. Results

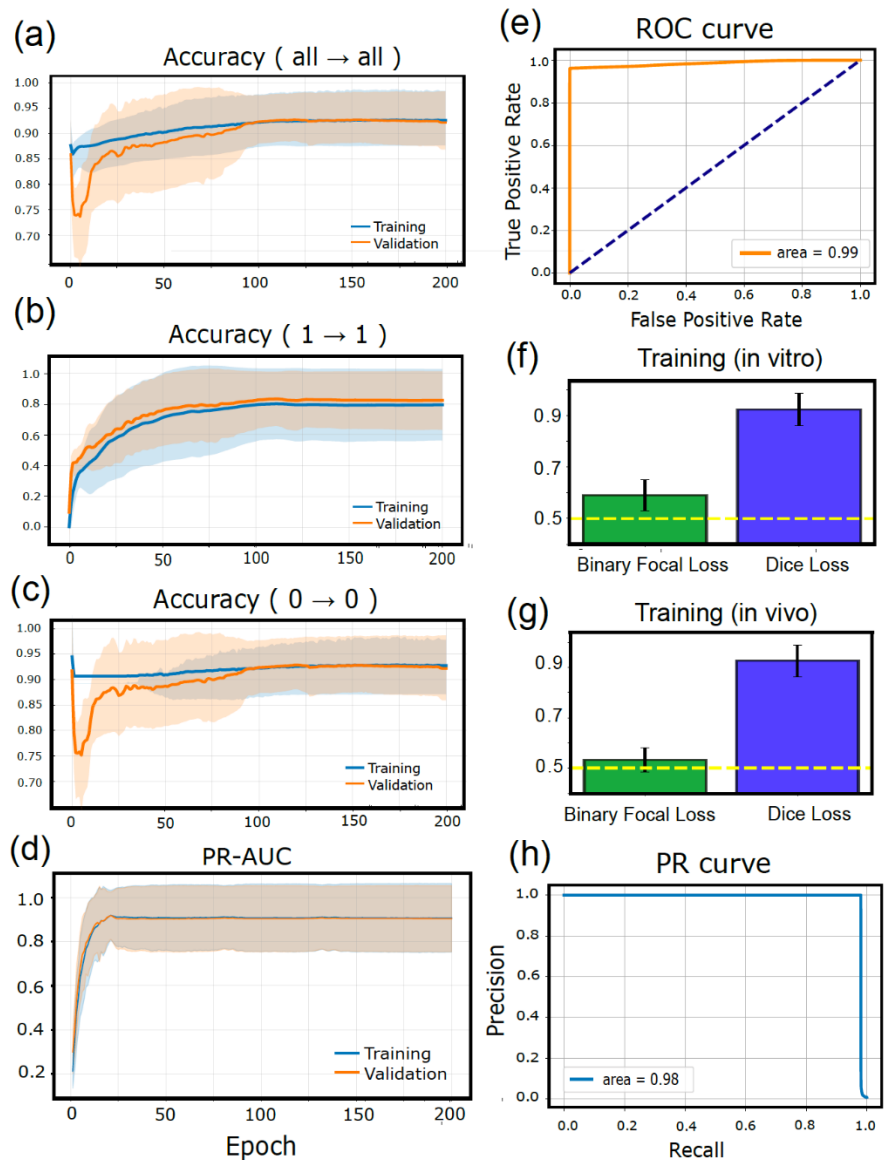
### 3.1. Evaluation and Comparison of Loss Functions During Training

In this study, we trained our model using in vitro data measured from slices of seven cortical regions in the left hemisphere, as well as six in vivo datasets recorded from either the cortex or hippocampus of the left hemisphere. To optimize the learning process, we employed Dice Loss as the loss function. As a result, across all training datasets, the area under the ROC curve (AUC) reached  $0.92 \pm 0.06$  (Figure 3). Because spike prediction is extremely class-imbalanced, we report ROC-AUC alongside Precision-Recall curves and PR-AUC, and we interpret PR-based metrics as the primary indicators of minority-event (spike) quality under extreme sparsity; ROC-AUC is retained to summarize ranking performance. During training, the true-positive and true-negative rates approached ~90% on average, but these values should be interpreted in the context of ROC/PR metrics. This high level of accuracy was difficult to achieve with alternative loss functions beyond the use of Dice Loss. Moreover, the spike (1) accuracy and PR-AUC also reached very high levels, supporting that minority-event (spike) detection improved beyond what ROC-AUC alone indicates.

Neural spike data present a unique challenge due to the extremely low frequency of the "1" state (spiking events). Consequently, predicting the occurrence of these essential "1" states is highly difficult. By employing Dice Loss as the error function, we effectively corrected the imbalance between the occurrence frequencies of 0 and 1, demonstrating its significant utility in this context (Figure 3a,b, c).

To further assess the effectiveness of Dice Loss, we compared its performance with the commonly used Binary Focal Cross Entropy Loss (Figure 3). In this study, we quantitatively evaluated predictive performance using the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. With Binary Focal Cross Entropy Loss, statistical significance was obtained only in self-prediction tasks using in vitro data. However, when Dice Loss was applied, the model achieved AUC values exceeding 0.9, demonstrating high prediction accuracy not only for in vitro data but also for in vivo data (Figure 3e,f).

These results indicate that Dice Loss effectively overcomes the challenge posed by the frequency imbalance between 0 and 1, significantly improving learning performance and achieving high-precision training outcomes.



**Figure 3.** Learning Process and Training Results.

Panel (a) illustrates the progression of accuracy over 200 training epochs across all datasets. Panel (b) shows how accuracy evolves for correctly predicting non-spiking (0) states during training, and panel (c) presents the corresponding accuracy changes for spiking (1) states.

Panel (d) reports the averaged PR-AUC over the 200-epoch training period across all datasets. To account for the class-imbalanced nature of spike prediction, we additionally report Precision-Recall curves and PR-AUC. Panels (e) and (h) present representative ROC and PR curves after training, respectively. Panels (f) and (g) compare performance across all samples for in vitro to in vitro and in vivo to in vivo prediction at the end of training, contrasting the binary focal loss and Dice loss; the numerical results are summarized in Table 2.

Taken together, the results for spike (1) accuracy and PR-AUC indicate that Dice loss improves minority-event (spike) detection beyond what ROC-AUC alone can capture.

### 3.2. Results of Generation

Our computational experiments demonstrated that training with Dice Loss improved within-domain and cross-domain prediction/generation. In the revised manuscript, we also include (i) training/validation loss curves for each loss function, (ii) depth/compute ablations, (iii) hyperparameter sensitivity analyses, and (iv) an explicit data-split table confirming independence of

animals/slices for in vivo to in vitro transfer. Representative time-series examples are shown in Figures 4c and 4d.

Figures 1e,f, and 3a illustrate progressive learning and strong generalization across datasets. Even in the lowest-scoring case, the ROC-AUC for in vitro and in vivo data generation was  $0.70 \pm 0.09$ , as summarized in Table 2 and Figure 4b.

**Table 2.** Performance Comparison Across Training Conditions.

| Conditions                        | Performance     |
|-----------------------------------|-----------------|
| In vitro (training)               | $0.92 \pm 0.06$ |
| In vivo (training)                | $0.93 \pm 0.06$ |
| In vitro 2 in vitro ( self data ) | $0.93 \pm 0.05$ |
| In vitro 2 in vitro ( others )    | $0.75 \pm 0.07$ |
| In vivo 2 in vivo ( self data )   | $0.77 \pm 0.10$ |
| In vivo 2 in vivo ( others )      | $0.74 \pm 0.10$ |
| In vitro 2 in vivo                | $0.70 \pm 0.09$ |
| In vivo 2 in vitro                | $0.80 \pm 0.10$ |

This table presents the performance metrics (mean  $\pm$  standard deviation) under different training conditions using in vitro and in vivo data. Both in vitro and in vivo data showed high performance during training with their respective data ( $0.92 \pm 0.06$ ,  $0.93 \pm 0.06$ ). Performance tended to decrease when using data from different sources compared to using self-data. However, generation between different environments also achieved performance above chance ( $AUC > 0.5$ ) and above our baseline comparisons.

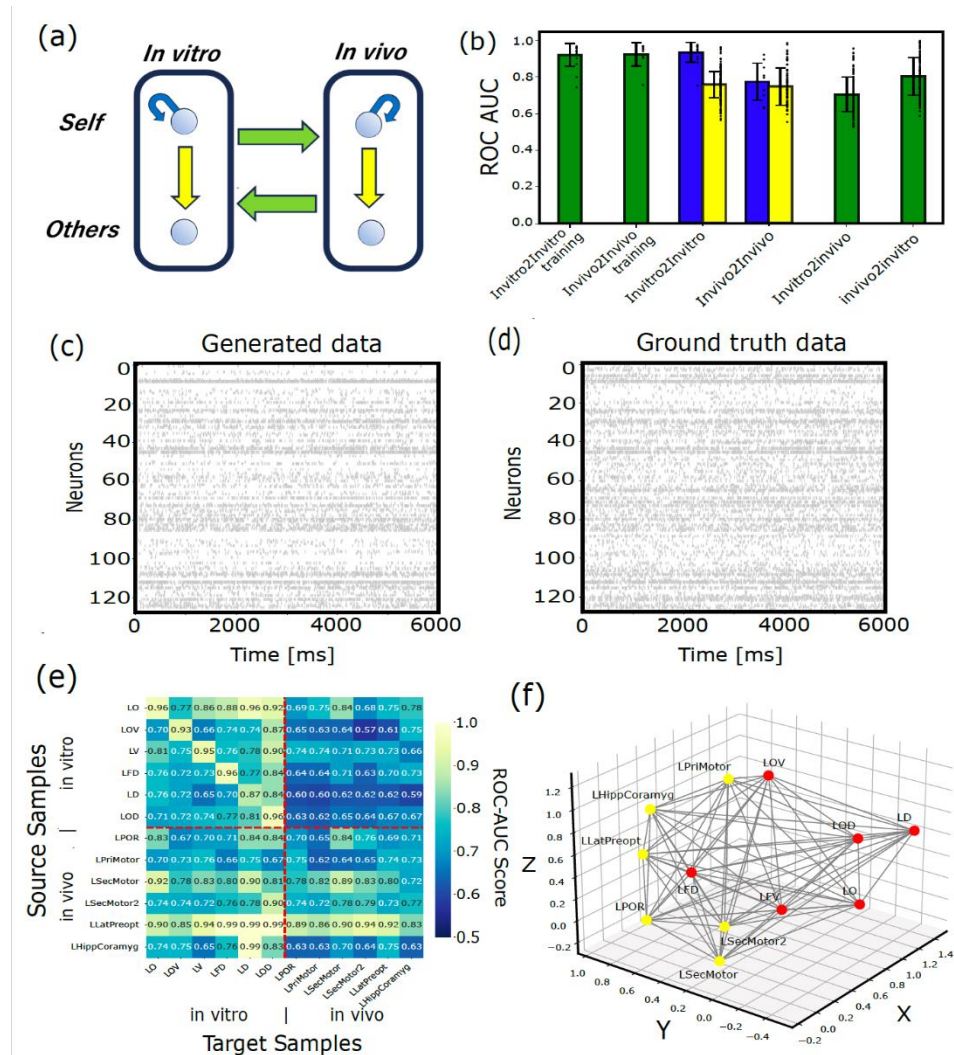


Figure 4. Results of Prediction-Generation Process.

Panel (a) presents a schematic diagram summarizing the various combinations of training and prediction data, effectively reframing Figures 1e,f. For *in vitro* data (circle at top left), both prediction of "Self" (curved blue arrow) between past and future own states and "Others" prediction (downward yellow arrow) to future states of other *in vitro* data are possible. Similarly, for *in vivo* data (circle at top right), both "Self" prediction-generation (curved blue arrow) and "Others" prediction-generation (downward yellow arrow) exist. Bidirectional generations between *in vitro* and *in vivo* data also exist, termed *in vitro*2*in vivo* and *in vivo*2*in vitro* for rightward and leftward generation, respectively.

Panel (b) displays bar graphs showing performance metrics for both training performance and prediction performances for the various combinations of training and prediction data outlined in (a). The leftmost two bars represent final training performance for *in vitro* and *in vivo* data. The next pair shows *in vitro*2*in vivo* and *in vivo*2*in vivo* generation, with each pair comprising "self" prediction (left bar) and "others" prediction (right bar). The rightmost two bars represent *in vitro*2*in vivo* and *in vivo*2*in vitro* generation.

Panels (c) and (d) present examples comparing generated time series with ground truth, with time on the horizontal axis and neuron index on the vertical axis.

Panel (e) displays a two-dimensional representation of performance metrics for combinations of training data (vertical axis) and prediction data (horizontal axis) across 12 datasets (6 *in vitro* and 6 *in vivo*). Red dotted lines divide the color map into quadrants: *in vitro* to *in vitro* generation (top left), *in vitro* to *in vivo* generation (top right), *in vivo* to *in vitro* generation (bottom left), and *in vivo* to *in vivo* generation (bottom right). The bar graph in (b) represents these grouped results, with ROC-AUC

scores on the vertical axis, separated into diagonal and non-diagonal components where applicable. Region abbreviations follow the IDs listed in Table 1.

Panel (f) represents the matrix from (e) reconfigured as a network diagram. It maps the data samples by optimizing positions after regarding the inverse of the AUC (used to measure prediction accuracy) as a distance metric in three-dimensional space. The red circular nodes represent in vitro data, while the yellow circular nodes represent in vivo data. The abbreviations for the region names follow the same conventions as in Table 1.

To analyze these results more comprehensively, we visualized the ROC-AUC scores for all combinations as a color map (Figure 4e). Figure 4f depicts the ROC-AUC color map matrix as a network diagram, where the inverse of the connection strength is treated as a distance and the positions are optimized in three-dimensional space. From these visualizations (Figure 4e,f), we identified five key findings:

First, in vitro to in vitro generation showed favorable performance with ROC-AUC scores around 0.93 between identical regions. While this exceptionally high performance was unexpected, the relatively strong performance in this combination was anticipated.

Second, interestingly, in vivo to in vivo generation did not demonstrate particularly superior performance between identical regions. We did not find evidence for greater non-stationarity in the in vivo data based on an Augmented Dickey-Fuller test ( $p=0.0083$ ); however, this result depends on our preprocessing choices and the time windows analyzed [24]. Our findings suggest that in vivo data exhibits stronger inter-regional influences compared to in vitro data, leading to variations in spike patterns.

Third, in vivo to in vitro generation outperformed in vitro to in vivo generation. This suggests that in vivo data encompasses the activity repertoire of in vitro data while partially acquiring new activity patterns. However, bidirectional comparisons indicate a non-trivial overlap with in vitro data in our setting.

Fourth, the lateral preoptic area (LPO) data showed relatively strong cross-region predictability within our sample. This finding will be extensively discussed in the Discussion section. In contrast to the lateral preoptic area, the cerebellum was less effective as a seed and more readily generated from other data in our sample; this may reflect simpler patterns in these recordings, but broader data would be needed to generalize.

Fifth, we could observe that in vitro data tends to cluster together with other in vitro data. At the same time, regions related to the cortical motor area, regardless of whether they are in vitro or in vivo, are concentrated in the central part. These characteristics support the idea that the ROC-AUC-based mapping meaningfully arranges the diversity of activity. Further insights can be gained by comparing this with Figures 1c and 1-d.

Sensitivity analyses and depth ablation (input length, depth, dropout, and number of heads) are reported in the Supplementary Materials (Text S1; Supplementary Figure S1 (validation) and Supplementary Figure S2 (generation)).

### 3.3. Analysis of Information Learned by the Model

To gain insight into the predictive mechanisms of the Transformer model, we conducted an analysis of the internal processing of information within the model. Here, to maintain focus and avoid unnecessary complexity, we limited our analysis to the case of in vitro to in vivo predictions.

In translation tasks using language models, it is well known that the correspondence between input words (before translation) and output words (after translation) is captured in the Attention Map [25,26]. Therefore, we began our analysis by examining the Attention Map (see the "Attention Map" subsection in the Methods section). Since our study deals with binary sequences (0s and 1s), we analyzed the relationship between the firing rate of the output signals and the weighted Attention Map. Specifically, to investigate how past information influences predictions, we examined how the average weight changes relative to the diagonal components of the Attention Map, which indicate time shifts from the present moment (Figure 5a). The results revealed a clear peak along the diagonal

component, suggesting that the model heavily relies on data from immediately preceding time points for its predictions. However, off-diagonal components were also observed, indicating that the model may be assigning supplementary attention to specific past moments.

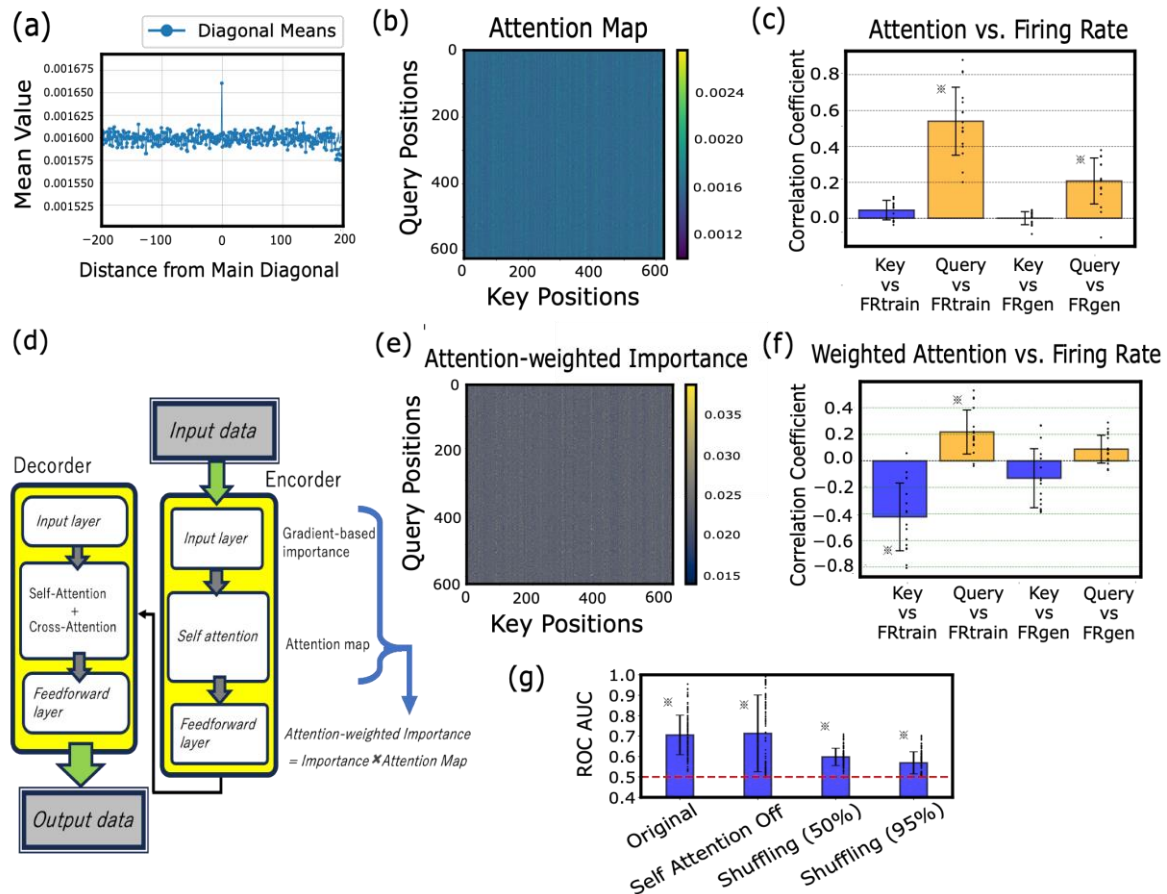
Next, we compared the Attention Map with the firing rates of input signals used during training. The results showed that the Attention Map correlated significantly with the query-side firing rate of the training data, whereas no significant correlation was found with the key-side firing rate (Figure 5b,c). To assess the critical role of the Attention mechanism in this learning process, we analyzed performance changes when the Attention mechanism was disabled. However, the results indicated that learning performance did not significantly deteriorate (Figure 5g, left two bars).

Given this outcome, we expanded our analysis beyond the Attention Map and introduced an importance measure based on gradient information, referred to as Attention-weighted importance (see the "Gradient-based Importance and Attention-weighted Importance" subsection in Methods). Specifically, we compared Attention-weighted importance with the firing rate of input signals during training. The results showed that while Attention-weighted importance exhibited a significant positive correlation with the query-side training data, it demonstrated a significant negative correlation with the key-side training data. Furthermore, generated data showed no significant correlation with either the query or key sides (Figure 5e,f).

However, in Attention-weighted importance, the Key axis (columns) primarily holds information related to high-firing regions of the input signals, and the attention weights are determined based on how the Query side references this information. On the other hand, the Query axis (rows) is influenced by importance-based weighting, and the distribution of attention is readjusted due to the effect of the loss function. Previously, the model focused primarily on high-firing regions, but it has been adjusted to also pay attention to low-firing regions (see Methods).

In fact, modifying the loss function to the Dice function significantly improved the model's predictive performance (Figure 3e,f). Within Attention-weighted importance, the Query side prioritized referencing distinctive information from high-firing regions of the input signals during training, primarily utilizing current and immediately preceding information. However, due to the influence of the Dice function's evaluation, the Query side's tendency in referencing Key-side information changed, and it was adjusted to also direct attention to distinctive features in past low-firing regions. As a result, the interaction between the priority derived from the input signals during training and the adjustment by the loss function contributed to the improvement in the model's predictive performance.

Finally, we investigated the extent to which prediction performance deteriorated when the input data used for prediction was shuffled across different cells within the same time window (Figure 5g, right two bars). As a result, when shuffling was applied across cells, the model's performance gradually declined as the window size expanded further into the past. However, even when 95% of the data was shuffled, the model retained a significant level of predictive accuracy (Figure 5g, rightmost bar). These results demonstrate that the model can still make reliable predictions even when the input data available for prediction is very limited.



**Figure 5.** Analysis of Information Learned by the Transformer Model.

Panel (a) illustrates representative attention maps within the Transformer model.

Panel (b) presents examples of attention-weighted importance maps, which are attention maps weighted by importance scores.

Panel (c) shows a schematic diagram depicting the relative relationships between input signals, attention maps, attention-weighted importance, and output signals.

Panel (d) demonstrates the intensity distribution of deviations from the diagonal components in the attention maps. A consistent peak near zero was observed across all datasets.

Panel (e) displays correlation plots comparing firing rates between training and generated data for both Key and Query sides of the attention maps. Results are shown as means with standard deviations, with individual data points plotted alongside error bars. Asterisks (\*) indicate statistical significance at  $p < 0.01$  (t-test).

Panel (f) presents correlation plots comparing firing rates between training and generated data for both Key and Query sides of the attention-weighted importance maps. The format follows that of panel (e).

Panel (g) displays four bar graphs. From left to right, they represent the following conditions: the original in vitro to in vivo generation case (second bar from the right in Figure 4b), a case where the attention mechanism is disabled, and prediction performance when the input data was shuffled across cells within each time slice for either 50% or 95% of the temporal window, extending from future time points to past ones. Asterisks (\*) indicate statistical significance at  $p < 0.01$  (t-test) when compared against the zero-correlation line.

## 4. Discussion

The primary contribution of this paper is algorithmic: predictive generation of spike trains under preparation/domain shift. As a practical implication, we highlight the value of this framework for in

vitro to in vivo transfer, where bridging across preparations remains challenging in neuroscience. We also emphasize that robust, sparsity-aware evaluation is essential for downstream applications, and we therefore report ROC-based metrics alongside Precision–Recall curves and PR-AUC.

Here, we discuss the key findings of this study, categorized into technical advancements in methodology and neuroscientific insights.

#### 4.1. Technical Advancements: The Role of Loss Function and Transformer Model

As mentioned in Section 2.2.2, the primary model used a single Transformer encoder layer. This design was chosen to retain a minimal-compute baseline and to reduce the risk of over-parameterization given the limited training data and the extreme sparsity of spikes. In the validation depth ablation (1–4 layers) under matched parameter budgets, the best depth varied by dataset and we did not observe a monotonic improvement with depth; depth = 1 remained competitive while deeper models increased compute. We therefore keep depth = 1 as the primary setting and present the full depth sweep in the Supplementary Materials.

To our knowledge, this is the first study to combine these two elements specifically for the generation of neuronal spike trains. Despite the simplicity of the idea, no prior work recognized the unique synergy between the Transformer architecture and Dice Loss in addressing the extreme class imbalance inherent to neural spike data. As shown in Figure 3e,f, this combination dramatically enhances predictive performance, outperforming the specific baselines we tested (Binary Focal Cross-Entropy and focal loss) on these datasets (Figure 3e,f).

It enables not only accurate within-modality generation (in vitro to invitro, in vivo to invivo), but also robust, bidirectional cross-domain generation between in vitro and in vivo conditions.

To understand the mechanisms underlying this high-accuracy generation, we first analyzed the core component of the Transformer model—the Attention Map. Our analysis revealed that the diagonal components of the Attention Map contributed significantly to model performance and that the query-side axis of the Attention Map exhibited a positive correlation with the firing rate of the input signals during training. To address a comprehensive interpretation of the model's behavior beyond Attention Map, we also introduced a gradient-based importance measure, weighting the Attention Map with importance scores, referred to as Attention-weighted importance. A comparison between Attention-weighted importance and the firing rate of input signals revealed that while the query axis showed a significant positive correlation, the key axis exhibited a significant negative correlation.

Importance strongly reflects gradient information and is highly sensitive to the choice of loss function. This suggests that the performance improvement associated with loss function selection primarily influences the region between the input layer and the self-attention mechanism in the Transformer model. This adjustment enables the model to incorporate long-term historical information, facilitating more refined learning rather than relying solely on firing rates.

While the results suggested a limited role for the Attention Map in predictive accuracy, it is important to acknowledge its potential contributions to training stability and learning speed.

#### 4.2. Neuroscientific Insights: Distinctive Brain Regions

Despite achieving high-accuracy predictions and generation, certain data exhibited particularly noteworthy characteristics.

The first notable finding concerns the meaningful multi-region mapping shown in Figure 4f. This mapping clearly captures the expected spatial relationships among data points in multiple aspects. For instance, the two data points measured from the secondary motor cortex are closely aligned and surrounded by regions associated with the motor cortex both of in vitro or in vivo data. Additionally, the in vivo and in vitro data are spatially separated into two distinct clusters on the left and right. These semantically meaningful embeddings represent the relative relationships between data points, suggesting how activity transitions from one data point to another. In this study, we demonstrate cross-generation from brief spontaneous activity in both traditional in vitro data and the

in vivo data provided by the International Brain Laboratory. Choosing the optimal embedding dimensionality is always a challenging problem. If one spatially maps neural-activity similarity by directly comparing it to actual spatial distances, both in vitro and in vivo points would naturally lie in three dimensional space. Hence, embedding the two datasets (in vitro and in vivo) in three to four dimensions is reasonably justified for their comparison in this work. However, should the number of datasets under comparison grow to three, four, five, or more, new justification will be required to determine whether a three dimensional visualization remains appropriate.

The second region of interest is the lateral pre-optic area (LPO) in vivo. This region appeared to be among the stronger seeds for predicting other brain regions in our dataset. Please note that the abbreviation "LPOR" used in Figure 4 and Table 1 refers to the left postrhinal area, which is a different region. The abbreviation for the lateral preoptic area in this context is "LLatPreopt".

The LPO, a hypothalamic nucleus, is one of the most extensively connected subdivisions of the hypothalamus: in the mouse it projects to and receives input from over 200 gray-matter regions, with intra-hypothalamic connections being especially prominent [27]. For example, it sends output to the lateral habenula (LHb), septal nuclei, ventral tegmental area (VTA), dorsal raphe nucleus, and the rostromedialtegmental nucleus (RMTg) [28]. It also receives input from the lateral septal nucleus (LS), which itself is a cortical hub region receiving hippocampal projections. Moreover, the LPO has reciprocal connections with the subiculum as well as CA1 and CA3 fields of the hippocampus.

The LPO is involved in both reward processing and the regulation of sleep-wake states [29,30]. Since sleep and wakefulness are fundamental behavioral states that entail wide-ranging shifts in brain function, the LPO's connectivity to arousal centers is critical. Recent work has also shown functional coupling between the LPO and the reward system: stimulation of the LPO suppresses GABAergic neurons in the VTA while increasing the firing rate of dopaminergic neurons [28]. Intriguingly, the LPO participates in an integrative circuit linking the nucleus accumbens—the ventral striatum involved in "positive" reward valuation—to the lateral habenula, which encodes aversive reward prediction errors.

Taken together, the LPO receives cognitive and sensory inputs from cortex and hippocampus and, via its outputs to brainstem and forebrain arousal and reward centers, functions as a hub that regulates global brain-state variables such as sleep-wake cycles, motivational drive, and homeostasis. This extensive functional connectivity allows LPO activity to "mirror" patterns of activity across diverse regions and, when necessary, influence network-wide dynamics. In other words, by virtue of its broad inputs, the LPO may internally represent spontaneous cortical activity patterns in miniature and emit potent control signals tightly linked to cortical dynamics. It is noteworthy that our data-driven analysis highlighted the LPO as a region of interest.

#### 4.3. Future Challenges: Expanding the Range of Applications

Based on these findings, two efficient strategies can be proposed.

First, the "proximity map" based on relative similarities between datasets expressed as a network diagram in Figure 4f is very important. This is because the proximity map encodes relative relationships among datasets, when a measurement is sparse or unavailable the framework can propose concrete substitutes—for example, selecting nearby datasets, mixing closely related datasets, or using the generative model to translate between conditions along the map's geometry. These are approximations of missing conditions rather than de novo creation, and they require task-specific validation.

From the standpoint of the 3Rs, this capability is a foundational technology that leads to the reduction of redundant experiments when existing results have been independently reproduced, while also prioritizing follow-up studies. That said, decisions to forgo new experiments should be based on predefined criteria and ethical review, and this method can be regarded as a technology that provides evidence for deliberations in that ethical review. In the future, as the number of nodes (datasets) in the network increases and network density grows within the informatics framework, the accuracy of generating non-existent data will steadily improve.

Second, prioritizing LPO measurements before expanding to other brain regions may enhance predictive accuracy and experimental efficiency because the Lateral preoptic area data provides good seeds for generating neural activity across many regions. This strategy aligns with the 3R principle (Replacement, Reduction, Refinement) in animal research and could improve efficiency in human neurophysiological studies. The complete explanation of why this region's neural activity can serve as such a versatile learning data seed (independent of the aforementioned proximity) remains unclear. As our understanding deepens, generation without requiring target data is expected to become increasingly feasible.

In the future, when considering the contribution of the LPO, some researchers in the life sciences may envision experiments involving optogenetic stimulation of this region. However, what is truly essential is to elucidate the "codes" utilized in the process of generating neural activity. In other words, it is crucial to uncover the information acquired by artificial neural networks through learning. To deepen our understanding of such phenomena, we expanded the interpretational scope of the Transformer's internal structure from attention maps to attention-weighted importance. Future challenges include further expanding this analysis and extracting and conducting detailed analysis of features from attention maps and Importance that contribute to prediction-generation.

In addition, it is important to improve methods for enhanced accuracy. Simple improvements include adding Position encoding to the Transformer model. As the computational method itself is scalable, expanding computational resources, such as computer memory, to increase the analyzable number of cells and time duration is also an important direction. While we performed mutual generation based on spontaneous activity, extending this to generate in vivo brain activity during stimulus presentation is another crucial direction. This can naturally be pursued by inputting in vivo spontaneous activity and stimulus information into a multi-modal AI model.

## 5. Conclusions

In conclusion, this study demonstrates a practical, high-precision method for generating neural activity and introduces a novel framework for dynamic integration and comparison of existing experimental datasets. By mapping spike-train dynamics rather than static connectivity, our approach may help reduce redundant replications of past experiments ("Reduction" of the 3Rs) and is intended to complement—rather than replaces—new experimental work. We will continue to improve accuracy, extend prediction horizons, and deepen our mechanistic understanding. In the future, this work may provide a basis for comparing analyses across animal and human datasets, but further validation will be required.

**Author Contributions:** MS was responsible for the study design, data-sharing negotiations, data analysis, and manuscript preparation.

**Funding:** MS was supported by multiple grants from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT; 21H01352, 23K18493).

**Institutional Review Board Statement:** Ethical approval. All animal procedures complied with institutional and national regulations. For in-vitro experiments, the use of C57BL/6J mice (3–5 weeks old) was approved by the Kyoto University Animal Experimentation Committee and performed in accordance with Kyoto University guidelines. A total of six mice were used in this study. The total number of mice used for the in vitro experiments in this study is six. For the in-vivo datasets analyzed here, all procedures were conducted in accordance with local laws and with approvals from the Animal Welfare Ethical Review Body of University College London; the Institutional Animal Care and Use Committees of Cold Spring Harbor Laboratory, Princeton University, the University of California at Los Angeles, and the University of California at Berkeley; the University Animal Welfare Committee of New York University; and the Portuguese Veterinary General Board, as written in International Brain Laboratory et al. 2023, 2025. The total number of mice used for the in vivo experiments is also six.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All original and generated datasets used in this study, together with related figures, are publicly available at Mendeley Data ("GenerativeNeurosci\_TFDice", V1; doi: 10.17632/kf65cvmtbz.1). The analysis and generation code is available at [https://github.com/ShimonoMLab/GenerativeNeurosci\\_ML-TrDic](https://github.com/ShimonoMLab/GenerativeNeurosci_ML-TrDic). For third-party source datasets, please refer to the original repositories and citations listed in the Methods and References. Abbreviations ACSF, artificial cerebrospinal fluid; AUC, area under the curve; IBL, International Brain Laboratory; LPO, lateral preoptic area; ROC, receiver operating characteristic.

**Acknowledgments:** MS is deeply grateful to GaeleChapuis of the International Brain Laboratory (IBL) and to Dora Angelaki, Tom Mrcic-Flogel, and Sonja Hofer for generously sharing invaluable neuronal activity data recorded with Neuropixels probes.

**Conflicts of Interest:** The author declares no competing financial interests.

## References

1. Dimakou, M.; et al. Spontaneous Activity (review). *Neuron* 2025. <https://doi.org/10.1016/j.neuron.2025.02.009>
2. Raichle, M.E., The brain's default mode network. *Annual review of neuroscience*, 2015, 38.1: 433-447.
3. Hartmann, T.S.; et al. Spontaneous Activity/Variability ("Noise") Perspective. *PLOS* 2015.
4. Bock, D.D.; Lee, W.-C.A.; Kerlin, A.M.; Andermann, M.L.; Hood, G.; Wetzel, A.W.; Yurgenson, S.; Soucy, E.R.; Kim, H.S.; Reid, R.C. Network Anatomy and In Vivo Physiology of Visual Cortical Neurons. *Nature* 2011, 471, 177–182. <https://doi.org/10.1038/nature09802>
5. Ding, Z.; et al. Functional Connectomics Reveals a General Wiring Rule in Mouse Visual Cortex. *Nature* 2025, 640, 459–469. <https://doi.org/10.1038/s41586-025-08840-3> (Nature)
6. The MICrONS Consortium. Functional Connectomics Spanning Multiple Areas of Mouse Visual Cortex. *Nature* 2025, 640, 435–447. <https://doi.org/10.1038/s41586-025-08790-w>
7. Zamani Esfahlani, F., Faskowitz, J., Slack, J., Mišić, B., & Betzel, R. F. (2022). Local structure-function relationships in human brain networks across the lifespan. *Nature communications*, 13(1), 2053. <https://doi.org/10.1038/s41467-022-29770-y>
8. Kajiwara, M., et al. Inhibitory neurons exhibit high controlling ability in the cortical microconnectome. *PLOS Computational Biology*, 2021, 17.4: e1008846.
9. Lepperød, M.E.; Stöber, T.M.; Hafting, T.; Fyhn, M.; Kording, K.P. Inferring Causal Connectivity from Pairwise Recordings and Optogenetics. *PLoS Comput. Biol.* 2023, 19, e1011574. <https://doi.org/10.1371/journal.pcbi.1011574>
10. Opitz, A.; et al. Limitations of Ex Vivo Measurements for In Vivo Neuroscience. *Proc. Natl. Acad. Sci. USA* 2017.
11. Wei, Y.; et al. Associations between In Vitro, In Vivo and In Silico Cell Classes in Mouse Primary Visual Cortex. *Nat. Commun.* 2023.
12. Nakajima, R.; Shirakami, A.; Tsumura, H.; Matsuda, K.; Nakamura, E.; Shimono, M. Mutual Generation in Neuronal Activity across the Brain via Deep Neural Approach, and Its Network Interpretation. *Commun. Biol.* 2023, 6, 1105.
13. Vaswani, A.; Shazeer, N.; Parmar, N.; et al. Attention Is All You Need. In *Adv. Neural Inf. Process. Syst.*; 2017, 30.
14. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; et al. An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale. In *Proc. ICLR*; 2021; arXiv:2010.11929.
15. Le, T.; Shlizerman, E. STNDT: Modeling Neural Population Activity with Spatiotemporal Transformers. In *NeurIPS*; 2022; arXiv:2204.01597.
16. YE, Joel, et al. Neural data transformer 2: multi-context pretraining for neural spiking activity. *Advances in Neural Information Processing Systems*, 2023, 36: 80352-80374.
17. Dice, L.R. Measures of the Amount of Ecologic Association between Species. *Ecology* 1945, 26, 297–302.
18. International Brain Laboratory; Aguillon-Rodriguez, V.; Angelaki, D.; Bayer, H.M.; Bonacchi, N.; et al. Standardized and Reproducible Measurement of Decision-Making in Mice. *eLife* 2021, 10, e63711. <https://doi.org/10.7554/eLife.63711>

19. Jun, J.J.; Steinmetz, N.A.; Siegle, J.H.; et al. Fully Integrated Silicon Probes for High-Density Recording of Neural Activity. *Nature* 2017, 551, 232–236.
20. Boussard, J.; Varol, E.; Lee, H.D.; Dethle, N.; Paninski, L. Three-Dimensional Spike Localization and Improved Motion Correction for Neuropixels Recordings. In *Adv. Neural Inf. Process. Syst.*; 2021, 34, 22095–22105.
21. Liu, Y.; Lapata, M. Text Summarization with Pretrained Encoders. In *Proc. EMNLP-IJCNLP*; 2019.
22. Jain, S.; Wallace, B.C. Attention Is Not Explanation. In *Proc. NAACL-HLT*; 2019; pp. 3543–3556.
23. Serrano, S.; Smith, N.A. Is Attention Interpretable? In *Proc. ACL*; 2019; pp. 2931–2951.
24. Dickey, D.A.; Fuller, W.A. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica* 1981, 49, 1057–1072.
25. Bahdanau, D.; Cho, K.; Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. In *Proc. ICLR*; 2015.
26. Voita, E.; Talbot, D.; Moiseev, F.; Sennrich, R.; Titov, I. Analyzing Multi-Head Self-Attention: Specialized Heads Do the Heavy Lifting, the Rest Can Be Pruned. In *Proc. ACL*; 2019; pp. 5797–5808.
27. Hahn, J.D.; Gao, L.; Boesen, T.; Gou, L.; Hintiryan, H.; Dong, H.-W. Macroscale Connections of the Mouse Lateral Preoptic Area and Anterior Lateral Hypothalamic Area. *J. Comp. Neurol.* 2022, 530, 2254–2285.
28. Gordon-Fennell, A.G.; Will, R.G.; Ramachandra, V.; et al. The Lateral Preoptic Area: A Novel Regulator of Reward Seeking and Neuronal Activity in the Ventral Tegmental Area. *Front. Neurosci.* 2020, 13, 1433. <https://doi.org/10.3389/fnins.2019.01433>
29. Saper, C.B.; Scammell, T.E.; Lu, J. Hypothalamic Regulation of Sleep and Circadian Rhythms. *Nature* 2005, 437, 1257–1263.
30. González, J.A.; Iordanidou, P.; Strom, M.; Adamantidis, A.; Burdakov, D. Awake Dynamics and Brain-Wide Direct Inputs of Hypothalamic MCH and Orexin Networks. *Nat. Commun.* 2016, 7, 11395.
31. Matsuda, K., Shirakami, A., Nakajima, R., Akutsu, T., & Shimono, M. (2023). Whole-Brain Evaluation of Cortical Microconnectomes. *eneuro*, 10(10).

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.