

Region-based conversion of neural activity across sessions

Woohyun Eum

*Dept. of Electrical and Computer Engineering
University of Florida
Gainesville, US
woohyun.eum@ufl.edu*

Carlton Smith

*Dept. of Physics
Yale University
New Haven, US
carlton.smith@yale.edu*

Shreya Saxena

*Dept. of Electrical and Computer Engineering
University of Florida
Gainesville, US
shreya.saxena@ufl.edu*

Abstract—A common way to advance our understanding of brain processing is to decode behavior from recorded neural signals. In order to study the neural correlates of learning a task, we would like to decode behavior across the entire timespan of learning, which can take multiple recording sessions across many days. However, decoding across sessions is hindered due to a high amount of session-to-session variability in neural recordings. Here, we propose utilizing multidimensional neural signals from Localized semi-non negative matrix factorization processing (LocaNMF) with high behavioral correlations across sessions, as well as a novel data augmentation method and region-based converter, to optimally align neural recordings. We apply our method to widefield calcium activity across many sessions while a mouse learns a decision-making task. We first decompose each session's neural activity into region-based spatial and temporal components that can reconstruct the data with high variance. Next, we perform data augmentation of the neural data to smooth the variability across trials. Finally, we design a region-based neural converter across sessions that transforms one session's neural signals into another while preserving its dimensionality. We test our approach by decoding the mouse's behavior in the decision-making task, and find that our method outperforms approaches that use purely anatomical information while analyzing neural activity across sessions. By preserving the high dimensionality in the neural data while converting neural activity across sessions, our method can be used towards further analyses of neural data across sessions and the neural correlates of learning.

Index Terms—Brain decoding, Calcium imaging, Decision making, Multi-session decoding

I. INTRODUCTION

Decision-making provides a rich experimental setting in which to explore the process of transforming a stimulus to a choice [1]–[4]. A two-alternative forced choice (2AFC) task is a common decision-making paradigm in mice, and allows for simultaneously recorded large-scale neural activity in head-fixed animals. However, the process by which an animal learns the 2AFC task has remained relatively under-explored. With the advent of large scale recording technologies that are able to record the neural activity with a high resolution over multiple sessions, we can now try to understand how learning of decision-making takes place across several days or even months. Widefield imaging calcium imaging (WFCI) provides an unprecedented view of neural data across the entire dorsal cortex of the mouse during these tasks. However, the

analysis of the resulting neural data across sessions poses several challenges.

Day-to-day noise and small shifts in the recording modality, such as imaging cameras, result in significant session-to-session variability in the recordings. Preprocessing methods include (a) pixel-wise averaging, resulting in one temporal signal per brain region, (b) singular value decomposition (SVD), resulting in components that cannot be directly assigned to any region of the brain, and (c) LocaNMF, resulting in a low dimensional set of temporal components anchored to each brain region. While pixel-wise averaging results in signals that are directly comparable across sessions, these signals do not capture the full extent of information in the brain region. SVD applied to the entire WFCI data makes it impossible to compare signals across sessions since the components may change drastically across sessions. Finally, while LocaNMF provides a lower-dimensional set of signals, it is still difficult to directly align the resulting components from one session to another. Alignment of neural data across sessions has been explored in the past for neural activity ranging from relatively low spatial resolution such as with functional Magnetic Resonance Imaging (fMRI), to single-cell neural recordings [5]–[7]. Here, we extend past methods by proposing a novel method to align WFCI data across sessions, and test out our algorithm on a rich WFCI dataset from 80 sessions across a mouse learning a 2AFC visual and somatosensory task.

In this paper, our main novelties are threefold. First, we present a new data augmentation method to train our models and unify the data across sessions. Second, we design a region-based neural converter to generalize temporal neural signals across sessions while preserving the dimensionality in the neural data across regions. Finally, we introduce a concrete test to evaluate our results: we decode the choice made by the animal using the converted data. Specifically, if we can decode the choice of the animal using the same decoder on converted data as trained on the original dataset, our conversion process can be considered a success. We compare our results with two established ways to align WFCI data across sessions. Using the results of the converter, we show how neural correlates evolve in each brain region during learning.

II. METHODS

A schematic overview of our method is shown in Figure 1. First, we preprocessed our data using a data augmentation method to decrease trial-to-trial variability during downstream steps. Subsequently, we designed a region-based neural converter across sessions to optimally align our temporal neural data across sessions. We also implemented two comparison methods for aligning data across sessions: clustering of spatial components, and pixel-wise averaging. These additional methods act as baselines for comparisons with the neural converter, and have been used before for analysis of WFCI [7]. Finally, we decoded the choice of the animal (left vs. right) using the converted neural signals as a measure of the performance of conversion. To quantify this, we calculated the accuracy of each model using the Area under the Receiver Operating Characteristic curve (AUC-ROC).

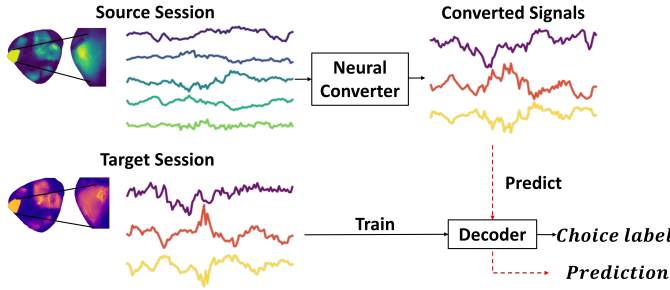


Fig. 1. A schematic overview of our methods: (a) data augmentation, (b) neural conversion of population activity, and (c) evaluation of converter using a choice decoder.

A. Experimental Methods

This experiment was conducted on a head-fixed mouse as it went through learning a visual and tactile delayed 2AFC spatial discrimination task over 80 sessions. The mouse was transgenic, expressing the calcium indicator GCaMP6f in excitatory neurons, and was imaged across multiple sessions using a custom-built widefield macroscope as in [8], to capture the entire dorsal cortex from above. The mouse was trained on two different modalities for the decision-making task: visual and tactile stimuli. The mouse licked the left or right spout to indicate choice after a delay period, and a water reward was provided if the choice direction was the same as the stimulus presentation side. To force the animal to commit to its initial decision, the opposite spout was moved out of reach after one spout was contacted. The mouse was observed over four months during 80 sessions as it performed experiments, and each session contained an average of 371 trials. The detailed experimental details are similar to those in [8].

B. Preprocessing

1) *LocaNMF Decomposition*: Calcium imaging is helpful for recording large neural populations, but may require downstream processing to recover the neural signals that we can use to understand the behavior. Moreover, it is difficult

to segment the data into signals pertaining to each region. Here, we perform temporal denoising, and decompose WFCI videos into temporal and spatial components using Localized semi-Nonnegative Matrix Factorization (LocaNMF) as in [9]. LocaNMF decomposition relies on the following mathematical modeling:

$$\hat{Y}(p, t) = \sum_k a_k(p) c_k(t) \quad (1)$$

Where $\hat{Y}$ is an estimated raw neural recording, a_k is a spatial component, c_k is a temporal component, p is pixel, t is time point, and $k \in [1, K]$, where K is the number of neural components. These spatial and temporal components are divided into different regions, and the temporal activity of any one region can now be represented using the following formula

$$X \sim f(tr, t, k) \quad (2)$$

where X is temporal activity of one brain region, tr is a trial, t is time point, k is the neural component, and f is the distribution of temporal neural activity from each session. We apply LocaNMF to each session's activity separately in order to model session-specific noise. Thus, we obtain 3-9 components per session in each brain region, representing the region's activity during the task. We would like to understand the relationship of the neural activity to the environment, and would thus need to align the neural activity across sessions.

2) *Data augmentation*: In order to have an adequate number of samples to train downstream models, we propose a data augmentation as in Figure 2. First, we divided entire trials into two groups depending direction of response (Left L or Right R). Second, we sampled a specified number of trials from each group. Third, we averaged over the sampled set and calculate the z-score. We refer to this as an 'augmented' trial. Finally, we assign each augmented trial with the corresponding label (L/R) depending on the group that it was sampled from. The advantage of this method is that we can generate as many augmented trials as we need. For this study, we generate an equal number of Left and Right augmented trials for the conversion and decoding process.

C. Conversion of Neural Activity across Sessions

To study how neural activity evolves during learning, we need to align the neural data across sessions and then relate the resulting neural correlates to behavior. For this goal, we introduce a neural converter that takes as inputs LocaNMF temporal components $\{c_k\}$ in one session, and outputs the temporal components in another session. This procedure is performed for each brain region, and is termed a 'region-based converter' here. We compare this converter to two established methods, spatial clustering and pixel-wise averaging.

1) *Region-based Converter*: The main task of the neural converter is transforming the source session's neural components to map to the target session's neural components, including the number of components. We designed our neural converter with a linear regression model:

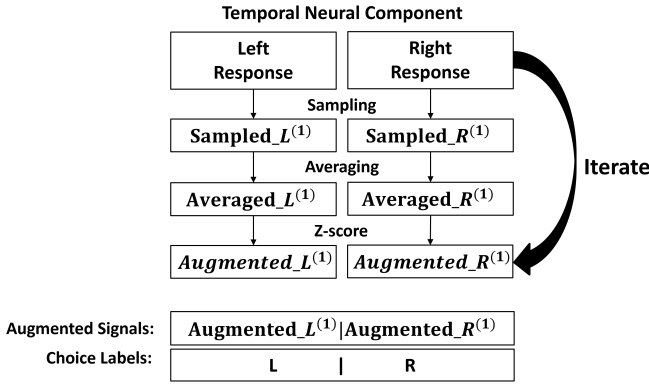


Fig. 2. Schematic for the data augmentation method. We sample trials with replacement according to the direction of response (L/R), and average the set. Next, we construct the labeled data set and iterate the process many times to get an adequate number of training and test samples.

$$\hat{Y}_i = W_{ij} X_j \quad (3)$$

Here, $\hat{Y}_i$ is the estimated target session's neural activity, here session i , with number of temporal components K_i . X_j is the neural signal of the source session, here session j , with number of temporal components K_j . Thus, W_{ij} is of dimension $K_i \times K_j$. Both X_i and Y_j are flattened matrices across times and trials. With our region-based converter, we are successfully able to recast the dimensionality of neural components from the source to the target sessions. After converting, we measured the R^2 score by comparing the estimated signals with the target session's signal to evaluate the performance of the converter.

2) *Spatial Clustering*: As a comparison to our region-based converter, we consider a method introduced in [7]. We cluster the spatial components resulting from LocaNMF in order to align the components, here using the identity of the neural populations that are activated across sessions. We first downsampled the images by 10% from the original size using 10×10 kernel size in order to reduce the dimensionality and increase the signal-to-noise ratio (SNR) of the high-dimensional spatial components. Next, we applied t-distributed stochastic neighbor embedding (t-SNE) to the downsampled dataset to further reduce the dimensionality of the spatial components. Based on the t-SNE map, we defined associated clusters using OPTICS, while setting a maximum euclidean distance between data points in each cluster. Among multiple clusters, the largest cluster was chosen as a correlated component, and we grouped temporal signals according to this cluster. The process of spatial clustering is shown in Figure 3.

3) *Pixel-wise average*: Using the brain atlas, here the Allen Atlas, we found pixels corresponding to each brain region. Next, we averaged the temporal signals from all the pixels in a given region, and this forms our one-dimensional pixel-wise average signal that can automatically be compared across sessions.

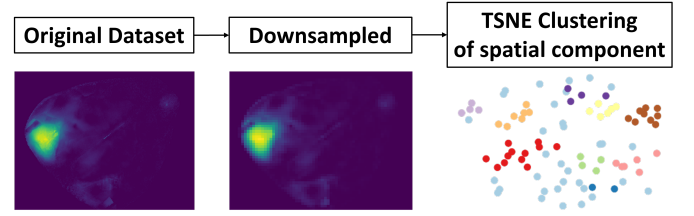


Fig. 3. Schematic and example of the spatial clustering method. To increase the SNR and reduce dimensionality, we downsampled the high-dimensional spatial components. Next, we applied t-SNE to further reduce the dimensionality of the data and to cluster the components. Using the resulting t-SNE map, we grouped the temporal neural components and performed decoding.

Using these two methods, spatial clustering and pixel-wise average, we built baselines for comparing the decoding performance of neural converter.

D. Decoding behavior across sessions

How do we know that our across-session neural converter is performing well? We test the performance of the converter by evaluating the ability of the converted neural data to decode the choice of the animal. We perform the following for data from every brain region separately.

We first train the decoder with the original data in a given session using a logistic regression [10]:

$$L^i = \sum_{j=1}^T \sum_{k=1}^K \frac{1}{1 + e^{-(\beta_0 + \beta_{jk} Y_{ijk})}} \quad (4)$$

Here, Y_{ijk} is the k^{th} component of the neural activity in the i^{th} trial and j^{th} time point in the trial, L^i is behavior label of i^{th} trial (L / R choice), and $\{\beta_0, \beta_{11}, \beta_{12}, \dots\}$ are estimated constants.

Next, we convert the session's data from all other sessions, i.e., we consider it as the 'target' session with each other session as the 'source' session. For each estimate of the session's data from other sessions, we then calculate the decoding accuracy while inputting the output of the converter (estimated 'target' session data) into the decoder. See Figure 1 for schematic. We perform this evaluation while considering each session as a target and all other sessions as sources. With the predicted labels, we quantify the decoding accuracy using the Area under the Receiver operating characteristics curve (AUC-ROC score) and calculate a confusion matrix to efficiently view the decoding results.

Since our label is composed of two classes (L/R) with an equal number of trials per class, our chance level is 50%. Therefore, we can validate our decoder by comparing the decoding accuracy with chance level, when it is trained and tested in the same session.

III. RESULTS

To understand how animal learns a decision-making task, we applied our converter - decoder methodology to data from 80 sessions of neural recordings. We evaluated the converter by measuring its decoding accuracy via the AUC-ROC score,

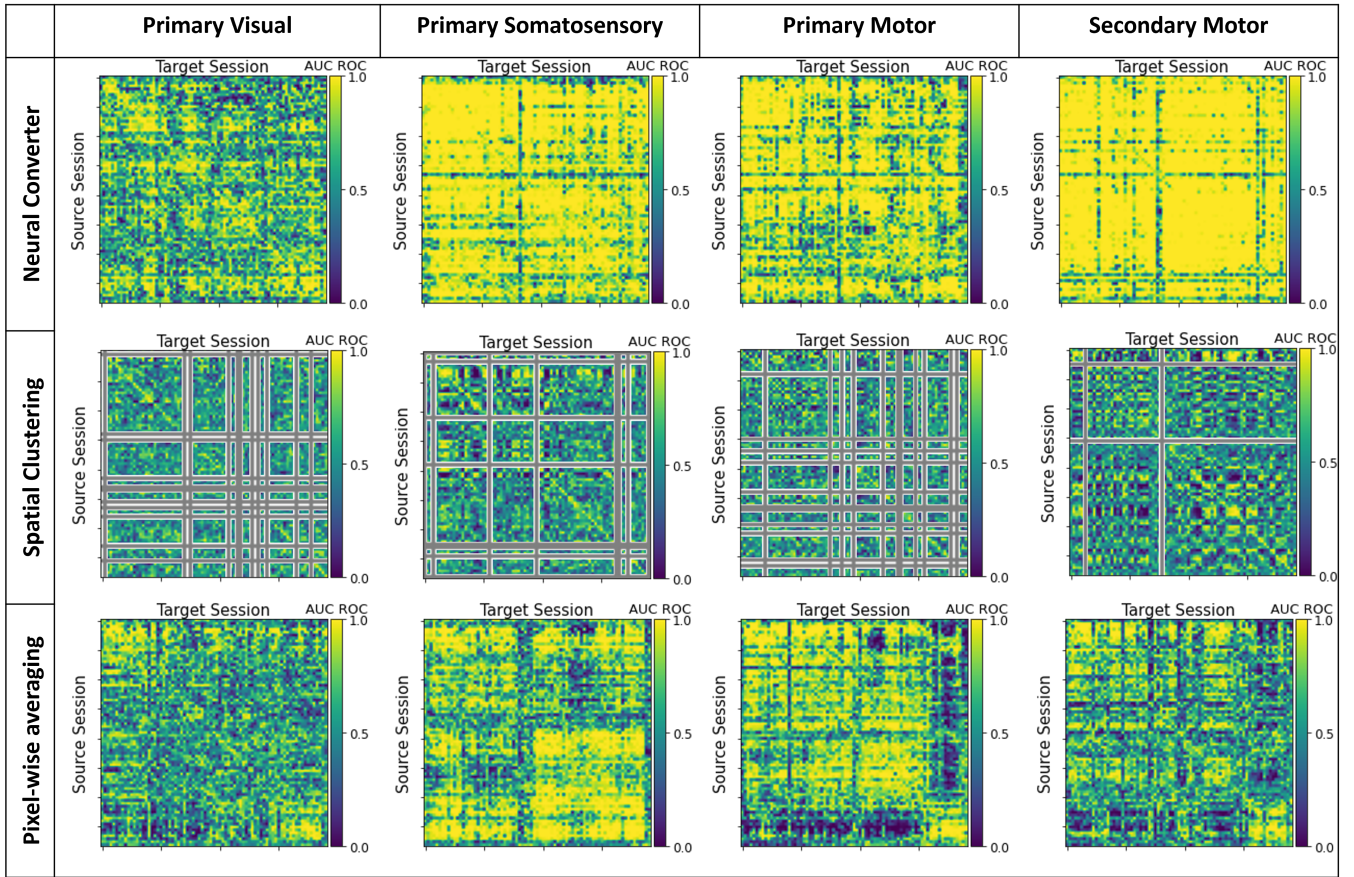


Fig. 4. Decoding accuracy of right side of brain regions: primary visual cortex, primary somatosensory cortex, primary motor cortex, and secondary motor cortex. Accuracy was measured through AUC-ROC score.

while converting the activity of each session to each other session. Moreover, we applied the two baseline methods to show comparisons of our conversion method with other more established methods.

The confusion matrix for the AUC-ROC is shown in Figure 4 for the three types of converter. In this matrix, the x and y axes represent the target and source sessions, respectively. We filled the diagonal points with the ‘upper bound’ values, which are the decoding scores from equally divided set for train and test in the same session of data. Since we used the same decoder on the converted signals as on the original dataset, each element of the confusion matrix conveys how well the neural signals were aligned between the source and target sessions using the three different conversion methods.

The confusion matrix for our proposed region-based converter attains a very high accuracy on most sessions, and vastly outperforms the baseline methods. The spatial clustering conversion method is not consistently able to convert the signals from each session to each other session. Lastly, the pixel-wise averaging works well, but is not able to achieve a high decoding accuracy. In fact, we note that the pixel-wise average does not always attain a high level of accuracy even while decoding the target session’s data itself (diagonal line). Thus, more dimensions from the original data may

be necessary in order to achieve a high accuracy, and our proposed region-based converter is able to use all dimensions in the neural data well.

We would have expected a block-diagonal structure in the decoding accuracy to convey that the neural data in neighboring sessions is somewhat stable, but here, the neural data seems to be convertible across all sessions equally well. There does not seem to be a noticeable chronological structure across in the decoding accuracy in any of the conversion methods.

We calculated the confusion matrix for the primary visual cortex, primary somatosensory cortex, primary motor cortex, and secondary motor cortex. Out of these, the primary visual did not convert across sessions as well as the other brain regions. This may be because the visual stimulus enters the brain in a different way than the tactile stimulus, and here we performed the same procedure on trials of all stimulus types.

IV. CONCLUSION

The study of session-to-session variability relies on the ability to be able to analyze the data from different sessions in the same space, to perform tasks such as decoding. Here, our region-based converter outperformed previously used methods for WFCI neural data. This converter consisted of a data augmentation step and then a linear conversion on extracted

region-based data. Our converter is able to act on multi-dimensional data to transform unordered signals from one session to another, and is able to decode the resulting choice of the animal with high accuracy. Future work includes the analysis of learning across the sessions using the learnt converters.

V. ACKNOWLEDGMENTS

We thank Anne Churchland for making the Widefield Calcium Imaging data available.

REFERENCES

- [1] Marios G. Philiastides, Roger Ratcliff, and Paul Sajda. Neural representation of task difficulty and decision making during perceptual categorization: A timing diagram. *The Journal of Neuroscience*, 26(35):8965, Aug 2006.
- [2] Roger Ratcliff, Marios G. Philiastides, and Paul Sajda. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the eeg. *Proceedings of the National Academy of Sciences*, 106(16):6539–6544, Apr 2009.
- [3] Marios G. Philiastides and Paul Sajda. Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral Cortex*, 16(4):509–518, Apr 2006.
- [4] Thomas Carlson, David A. Tovar, Arjen Alink, and Nikolaus Kriegeskorte. Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, 13(10):1–1, Aug 2013.
- [5] Juan A. Gallego, Matthew G. Perich, Raed H. Chowdhury, Sara A. Solla, and Lee E. Miller. Long-term stability of cortical population dynamics underlying consistent behavior. *Nature Neuroscience*, 23(2):260–270, Feb 2020.
- [6] Kentaro Yamada, Yoichi Miyawaki, and Yukiyasu Kamitani. Inter-subject neural code converter for visual image representation. *NeuroImage*, 113:289–297, 2015.
- [7] Simon Musall, Xiaonan R. Sun, Hemanth Mohan, Xu An, Steven Gluf, Shujing Li, Rhonda Drewes, Emma Cravo, Irene Lenzi, Chaoqun Yin, Björn M. Kampa, and Anne K. Churchland. Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. *bioRxiv*, 2022.
- [8] Simon Musall, Matthew T. Kaufman, Ashley L. Juavinett, Steven Gluf, and Anne K. Churchland. Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, 22(10):1677–1686, Oct 2019.
- [9] Shreya Saxena, Ian Kinsella, Simon Musall, Sharon H Kim, Jozsef Meszaros, David N Thibodeaux, Carla Kim, John Cunningham, Elizabeth MC Hillman, Anne Churchland, et al. Localized semi-nonnegative matrix factorization (locanmf) of widefield calcium imaging data. *PLoS computational biology*, 16(4):e1007791, 2020.
- [10] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2006.