Exploring prominent factors of behavioral correlations and the alignment to Brain Atlas

Eric Chen
Dept. of Electrical and Computer Engineering
University of Florida
Chen.Eric@ufl.edu

Woohyun Eum
Dept. of Electrical and Computer Engineering
University of Florida
woohyun.eum@ufl.edu

***Abstract*—The vast amount of previous animal behavior analysis revealed that most behavioral information is encoded in correlations of neurons. Therefore, capturing the correlations from neural recordings has become the primary workhorse to explore and understand the neural population activities, response structure, and functional connectivity of the brain. However, uncovering prominent factors determining the correlations is difficult to analyze and it changes in different experimental settings. In order to find the optimum strategy to find prominent features determining correlations, we aim to capture them from different latent spaces by using T-distributed Stochastic Neighbor Embedding. Hence, this proposed research focuses on comparing behavioral correlation with brain parcellation by using the Allen brain atlas as a reference for anatomical information.**

I. Introduction and Motivation

Recordings of neural population activities allow researchers to gain insights regarding various cognitive and motor functions that are based on complex neuron interactions in animal behavior analysis [1]–[3]. Common views in neuroscience research identify that neurons' behaviors are not likely to be independent of each other, thus, are governed by common underlying pattern(s) or cause(s) – *latent*. Previous studies introduced multiple machine learning (ML) approaches to unveil the latent neural population activities to form new hypotheses [3], to examine the population response structure [4], [5], and to segment the anatomical regions of brain atlas [2].

The selection of appropriate algorithms to disambiguate proper latent space may have a significant bearing on deriving interpretable and accurate conclusions in these neuro-encoding research areas [1]. Existing algorithms vary in the statistical methods that affect inherent loss and the preservation of the information from neural activities. Hence, in this proposed work, we will compare and evaluate how the commonly adopted ML algorithms – T-stochastic neighbor embedding (t-SNE) and Principal Component Analysis (PCA) – perform in the task of segmenting the anatomical regions of the brain atlas using the neural population recording datasets.

In this study, our goal is to explore the underlying features of behavioral correlates by classifying brain regions based on correlations, using neural population recording datasets. We believe that the application of commonly adopted machine learning (ML) algorithms can be useful for this purpose. In particular, we are interested in exploring both unsupervised and supervised methods, such as T-stochastic neighbor embedding (t-SNE) and Principal Component Analysis (PCA). By comparing and evaluating these algorithms, we hope to gain insights into how they perform in the task of segmenting the anatomical regions of the brain atlas.

II. Related Work

PCA and t-SNE (t-Distributed Stochastic Neighbor Embedding) are two commonly used data dimensionality reduction techniques used in the field of machine learning and data science. Both techniques can be used to understand neural correlation data in neuroscience research. PCA is a technique that is commonly used to reduce the number of dimensions in a dataset while retaining the most information by finding the correlation between dimensions [2]. On the other hand, t-SNE is a technique that is used to visualize high-dimensional data by reducing the number of dimensions while preserving the structure of the data as much as possible.

The t-SNE is a widely used algorithm to find clustered neural activities and visualize those correlations in two-dimensional space. t-SNE (t-Distributed Stochastic Neighbor Embedding) is a powerful data visualization tool used for dimensionality reduction and clustering analysis. It is widely used in various fields, including biology, computer science, and social sciences. t-SNE is particularly useful when visualizing high-dimensional data with complex relationships between variables. According to a study by [6], t-SNE outperforms other dimensionality reduction techniques such as PCA in preserving the local structure of the data.

When applied together, PCA can be used to preprocess the data and then t-SNE can be applied to reduce the number of dimensions further to provide a visualization of the data. This approach has been used in neuroscience research to understand neural correlation data. For example, a study used PCA to preprocess the neural correlation data from monkeys while they were performing a reaching task and then applied t-SNE to visualize the data in a two-dimensional space [7]. This approach allowed researchers to identify patterns and relationships between neurons that could not have been identified by analyzing the data in its original high-dimensional space.

III. Problem Statement

The purpose of this research is to compare and evaluate the application of commonly adopted machine learning algorithms, specifically T-stochastic neighbor embedding (t-SNE) and Principal Component Analysis (PCA), for segmenting the anatomical regions of the brain using neural population recording datasets. This study aims to explore both unsupervised and supervised methods for identifying the underlying prominent features of behavioral correlates. Our model will classify brain regions based on correlations, providing insights into the relationships between behavioral correlates and brain activity. Through this research, we seek to contribute to a deeper understanding of the neural mechanisms underlying behavior and cognition.

IV. Proposed Methods

A. Data Collection and Preparation

We will use the brain-wide recording dataset from [8]. The dataset contains recordings of over 30,000 neurons across 42 different mouse brain areas, when performing a visual discrimination task. Specifically, we will analyze the recordings of the two specific regions of visual cortex (i.e., VISa, VISp) and two regions of hippocampus (i.e. CA1, POST) of a single mouse in one sessions in the left-response trial.

B. Analysis Procedure

The analysis was conducted in three steps. We conducted PCA and t-SNE analyses in the Python 3.8. environment in order finding optimum embedding space algorithms. We chose these two algorithms because they vary in noise reduction, time sequencing, and temporal smoothing.

First step of the analysis is to implement a Principal Component Analysis (PCA) to reduce the dimension of the original dataset that comprised on the information about neuron firing rate across time and across a trial. Especially, we extracted the from the two specific visual cortex regions (i.e., VISA, VISPM) and hippocamal regions (i.e., POST, CA1). For instance, our dataset comprises a total of 914 neuron recordings across one trial for 250 ms. The PCA analysis was applied to reduce the data dimension w.r.t trials and binned time, resulting in the latent matrix $M^{K\times 62750}$, where K represents the number of underlying latent. We evaluated the range of final component sizes based on the total explained variance.

In the second step, we conducted a t-SNE algorithm using the results we obtained from the PCA analysis. The similarity across the two regions are evaluated using Kullback-Leibler (KL) divergence. The outcome of this analysis were clustered neural recordings of the four specific regions (i.e., VISA, VISPM, CA1, POST). Several variation versions of visualization were created by applying t-SNE to varying number of component sizes. Specifically, we evaluated the t-SNE results that were generated from PCA results of component sizes range from 50 to 300.

Last, we implemented several classification algorithms to provide the PCA results as input and provide classification on their brain region (i.e., VISA, VISPM, CA1, POST). Specifically, we compared the logistic regression, Random Forest classifier, and multi-perceptron for comparisons. The multi-class classification performance was compared based on varying measures, including accuracy and kappa score accumulated from the 5-fold cross-validation results. We reported the 5-fold cross-validation performance results on the training set (70% of the data), whereas the test performance was reported on the test set (30% of the data).

Further, we evaluated how the components that contributed to the high performance accuracy could be explained and interpreted using several common hand-engineered time-series features. The hand-engineered features were adopted from [9]. The Pearson's correlation coefficients were evaluated to explore the interpretability of the PCA component scores.

V. Results

A. Principal Component Analysis Results

The finding suggests that by analyzing up to 300 components retrieved from a dataset that was flattened to represent neurons across trials and time, it was possible to explain close to 80% of the original variance in the data. However, each individual component only accounted for a small percentage of the total variance, with the first four components accounting for 5.7%, 1.4%, 1.1%, and 1.0%, respectively. This implies that a large number of components are needed to capture the majority of the variance in the data. Hence, we subsequently evaluated how the use of varying component sizes within the range (0<k<100) could be Incorporated with t-SNE to visualize the sub-regions of visual cortex and hippocampus. Figure 1 provides the total explained variance in PCA based on the number of extracted components.

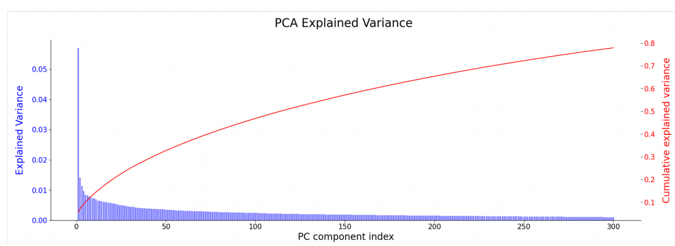

Fig. 1. PCA Results and the Explained Variance

B. PCA and t-SNE Visualization Results

Figure 2 provides an overview of the visualization results using t-SNE with the PCA results. The t-SNE visualization results varied depending on the number of PCA components evaluated. The results showed a fairly distinguishable segmentation between the different brain regions, but the effect of increasing the number of components on the visualization results was also observed. There is a trade-off effect between increasing the components and the t-SNE separation results. Some experts argue that including more components can make clusters more distinguishable and cluster-able, while others argue that including more components may introduce more noise to the data and make the separation task more challenging.

Hence, our next results address this issue by implementing and comparing classification models using supervised learning method to find the balance points between the number of components and the t-SNE separation results and achieve the best classification accuracy.

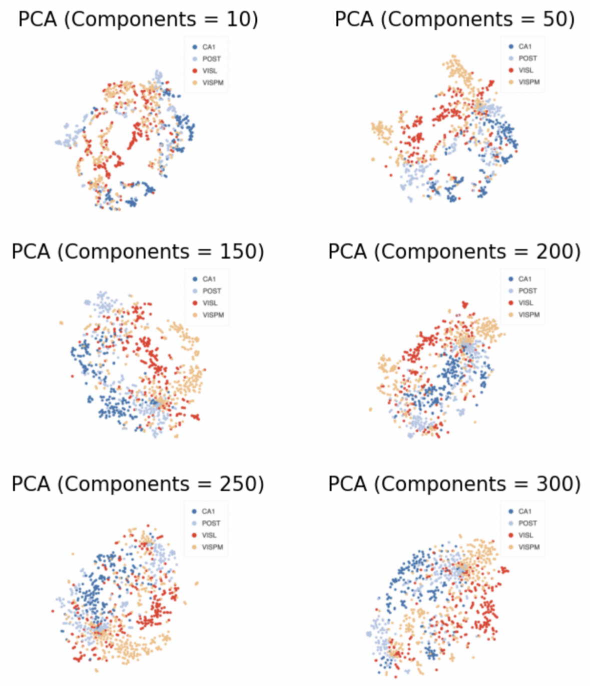

Fig. 2. PCA with T-SNE Visualization Results with varying components

C. PCA and Classification Model Performance Results

Tables 1 and 2 provides the summary of the classification performance on the train set. The RF and MLP classifiers performed comparably, achieving close to 80% accuracy. However, the accuracy varied depending on the brain region, with the highest accuracy observed in VISPM and the lowest in POST. The best performance was identified at number of components=60 for both MLP and RF, which is around 43% of the original variance explained in the train data. To evaluate the statistical significance of the accuracy comparisons, a confidence interval of 95% was used. If the confidence intervals overlap, it indicates that the two categories are not statistically different. The confidence intervals were based on the standard deviation of 5-fold cross-validation. On the other hand, the logistic regression model did not perform well. This could be attributed to a non-linear relationship between the components, which cannot be effectively handled by a linear model such as logistic regression. When plotting components against each other to evaluate the underlying cause, a non-linear relationship was observed, which could be why logistic regression performed poorly.

In Figure 3, the association between the number of PCA components and the classification accuracy of the Random Forest (RF) and Multi-Layer Perceptron (MLP) classifiers on the test set is shown. The maximum classification performance was achieved when n-components=170 and n-components=60, accounting for approximately 71% and 43% of the total variance for RF and MLP, respectively. The best accuracy reported on the test set was 81% for RF and 82% for MLP. The figure demonstrates that both classifiers exhibit a general pattern of increasing accuracy as the number of components increases up to 60 components, after which accuracy plateaus. Further increasing the number of components does not appear to significantly improve classifier performance.

TABLE I
MULTICLASS CLASSIFICAITON PERFORMANCE COMPARISONS

	Accuracy	Kappa	Accuracy (C.I.)
Logistic Regression	0.71	0.60	(0.70, 0.72)
Random Forest	0.80	0.73	(0.79, 0.81)
Multi-layer Perceptron	0.79	0.72	(0.78, 0.80)

TABLE II
CLASSIFICATION PERFORMANCE ACCURACY BY CLASS

	CA	POST	VISL	VISPM
Random Forest	0.80	0.73	0.81	0.80
Logistic Regression	0.77	0.71	0.74	0.90
Multi-layer Perceptron	0.77	0.77	0.79	0.91

D. Explaining the PCA components

The first 10 PCA components were evaluated based on their correlation to commonly used time-series feature sets. These feature sets were computed based on neuron x averaged over time. The first feature set is mean firing rate across trial time, which is the average number of times a neuron fires in response over a certain period of time and across trial.

The second feature set is mean firing rate entropy across trial, which is a measure of the variability or randomness in the average firing rate of a neuron during time time bean across trial. The third feature set is the number of crossings, which pertains to the frequency at which the firing rate of a neuron crosses a certain threshold or boundary (in this case, 6) in average over time. The fourth feature set is linear trend time wise averaged trial, which refers to the point at which the firing rate of a neuron or population of neurons intersects with the y-axis (i.e., the firing rate when time is equal to zero) in a linear regression analysis.

The fifth feature set is quadratic trend, which is $a + btime + ctime^2$. The last feature set is mean change across time averaged over trial and count above 100 total sum, which is the number of neurons that had more than the mean count of 100 across trials and average over time. These feature sets were used to evaluate the first 10 PCA components to determine their correlation to commonly used time-series features.

Specifically, Component 1 was strongly correlated with the mean firing rate across trial and time, as well as the mean firing rate entropy across trial. This indicates that component 1 is largely influenced by the variability and average firing rate of neurons over time and across trials. Additionally, component 1 was also found to be moderately associated with the CID-CE feature, which pertains to the variability of the firing rates

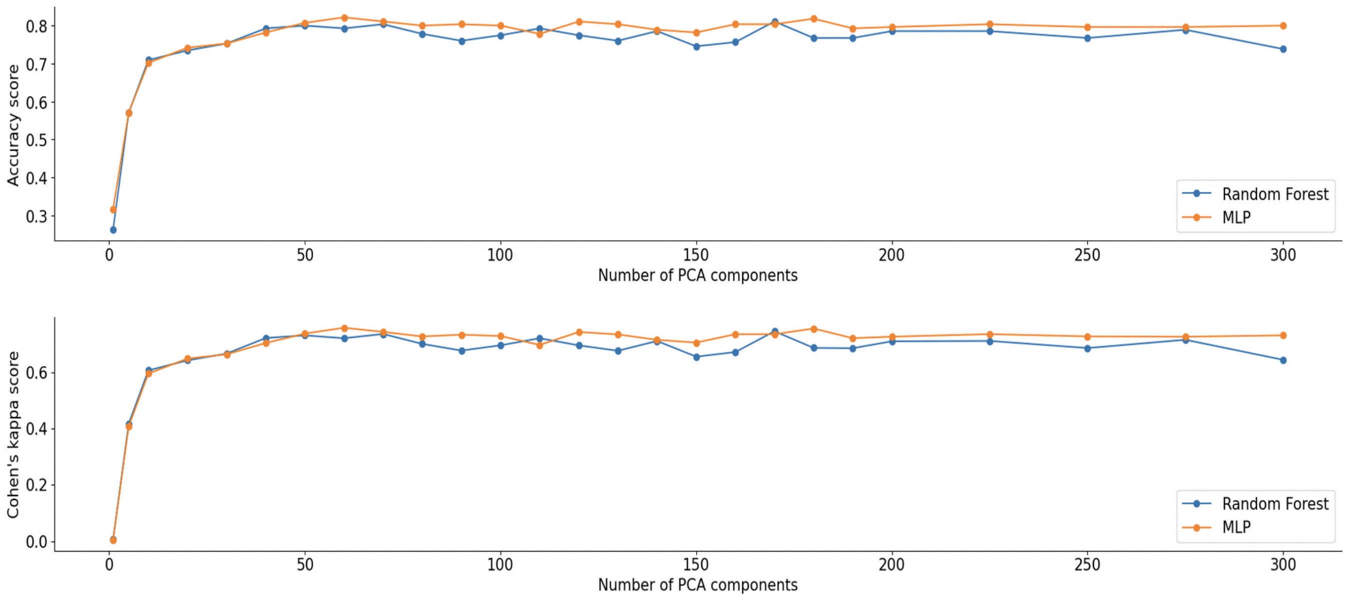

Fig. 3. Classification Performance based on the number of PCA components

of individual neurons in a population. These findings suggest that component 1 captures significant information about the average and variability of the neural activity in response to a certain task or stimulus.

Similarly, Component 2 showed a positive correlation with the linear-trend-timewise averaged trial feature set, indicating that it was associated with the intersection point of the firing rate of neurons with the y-axis in a linear regression analysis. On the other hand, component 2 had a negative correlation with the quadratic trend timeline feature set, which represents the firing rate of neurons changing in a nonlinear pattern over time. This suggests that component 2 was not strongly influenced by the quadratic trend timeline feature set and may be more influenced by the linear trend over time. Therefore, researchers should consider both the linear and quadratic trends when evaluating the influence of component 2 on neuronal firing rates.

VI. Conclusion

In conclusion, our study demonstrates that the correlation information of neuron population recordings can effectively segment brain regions. Both unsupervised and supervised methods were evaluated, with PCA and t-SNE exhibiting a fairly distinguishable pattern to cluster neurons based on regions. RF and MLP classifiers performed the best with around 80% accuracy, consistent with previous literature.

Our findings provide valuable insights into the understanding of how neuron populations in different brain regions interact and contribute to brain functions. Especially, by attempting to understand the correlation between hand-engineered time-series features with the components, we could discover the most promising features that contributed to our classification performance. Especially, the average firing rate of the neurons and the linear trends of firing rate across time were identified as promising features.

Future directions include expanding the interpretation of components with in-depth time-series analysis and evaluating the generalization of the current methods and findings to neuron recordings of other brain regions. There are several other avenues of research that could be explored based on the findings of this study. For example, it would be interesting to investigate the relationship between the identified components and specific behaviors or tasks performed by the animals. This could provide valuable insights into how different neural populations contribute to different cognitive or motor functions.

Furthermore, it may be possible to extend the current methods to analyze recordings from other brain regions and species, which could help to uncover general principles underlying neural coding across different contexts. Additionally, the methods used in this study could be modified and adapted to address other questions related to neural activity, such as the detection of neural oscillations or the identification of specific patterns of activity associated with particular diseases or disorders. By continuing to build on the findings of this study, we can gain a deeper understanding of the complex mechanisms that underlie brain function and behavior, which could ultimately lead to the development of new treatments for neurological and psychiatric disorders. Overall, this study contributes to the growing body of knowledge on decoding the neuronal code and its underlying mechanisms in the brain.

Author Contribution

Data Preparation and Cleaning: EC and WE; t-SNE Implementation: WE; PCA and Classifiers Implementation and Interpretation: EC; Presentation Preperation: EC and WE; Report Writing: EC and WE

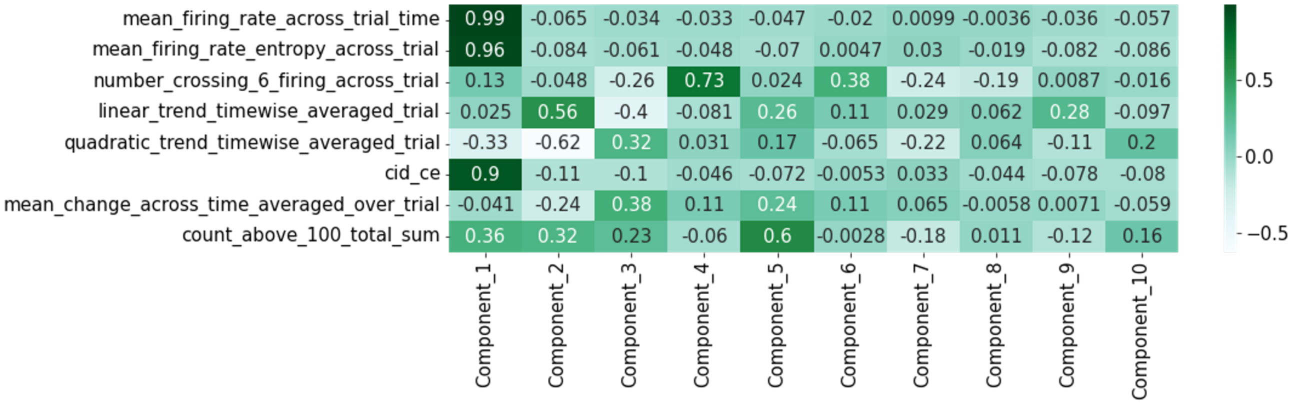

Fig. 4. PCA Components Correlation Results with the Hand-engineered Time-series Features

References

[1] J. P. Cunningham and B. M. Yu, "Dimensionality reduction for large-scale neural recordings," *Nature neuroscience*, vol. 17, no. 11, pp. 1500–1509, 2014.

[2] A. Iqbal, R. Khan, and T. Karayannis, "Developing a brain atlas through deep learning," *Nature Machine Intelligence*, vol. 1, no. 6, pp. 277–287, 2019.

[3] J. Freeman, N. Vladimirov, T. Kawashima, Y. Mu, N. J. Sofroniew, D. V. Bennett, J. Rosen, C.-T. Yang, L. L. Looger, and M. B. Ahrens, "Mapping brain activity at scale with cluster computing," *Nature methods*, vol. 11, no. 9, pp. 941–950, 2014.

[4] M. T. Kaufman, M. M. Churchland, S. I. Ryu, and K. V. Shenoy, "Cortical activity in the null space: permitting preparation without movement," *Nature neuroscience*, vol. 17, no. 3, pp. 440–448, 2014.

[5] O. Mazor and G. Laurent, "Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons," *Neuron*, vol. 48, no. 4, pp. 661–673, 2005.

[6] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne." *Journal of machine learning research*, vol. 9, no. 11, 2008.

[7] J. Pareek and J. Jacob, "Data compression and visualization using pca and t-sne," in *Advances in Information Communication Technology and Computing: Proceedings of AICTC 2019*. Springer, 2021, pp. 327–337.

[8] N. A. Steinmetz, P. Zatka-Haas, M. Carandini, and K. D. Harris, "Distributed coding of choice, action and engagement across the mouse brain," *Nature*, vol. 576, no. 7786, pp. 266–273, 2019.

[9] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package)," *Neurocomputing*, vol. 307, pp. 72–77, 2018.