BRAIN INFORMATION PROCESSING ANALYSIS USING ARTIFICIAL INTELLIGENCE METHODS

Bahareh Rahmani [1], Han Yiyuan [2], Muhammad Aqeel [1], Payam Norouzzadeh [1], Abbas Maazallahi [3], Eli Snir [2], Ghazaleh Ashrafi [2]

[1] Saint Louis University, School of Medicine, Saint Louis, MO, USA
[2] Washington University in Saint Louis, School of Business, Saint Louis, MO, USA
[3] University of Tehran, School of Engineering, Tehran, Iran

ABSTRACT

The human brain is a complex organ that processes millions of neurons and transmits information through billions of synapses. Here, we perform a quantitative analysis of a fluorescent reporter of synaptic vesicle release in synapses to gain insights into the underlying patterns of synaptic transmission. In these experiments, dissociated rat hippocampal neurons expressing the reporter were electrically stimulated with field potential, and fluorescence signals were recorded. We observed a positive correlation between the resting intensity level after stimulation and resting intensity level before stimulation, peak value, and time interval to peak. These findings provide valuable insights into the response of individual synapses to electrical stimulation and uncover important features of synaptic transmission.

KEYWORDS

Brain, Neurons, Synapses, Fluorescence Intensity, Electrical Stimulation, Statistical Analysis, Regression

1. INTRODUCTION

Neuroscience is undergoing rapid change with the emergence of large datasets on connectomics, synaptic activity, optogenetic manipulations of brain circuits, and behavioral assays. These developments necessitate the application of advanced statistical and analytical methods in neuroscience to infer meaningful knowledge from large datasets [1][2]. With growing data availability, information granularity, and analytical programming, data science has become a crucial tool in neuroscience. Data science has been used to develop new methods for analyzing and visualizing brain data, such as interactive visualizations and graph-based representations [3]. Moreover, it has been employed to analyze large-scale datasets, such as the Human Connectome Project, to better understand the structure and function of the brain [4]. Analytical techniques, including statistical analysis and machine learning, are also being used to identify patterns and relationships in synaptic activity [5].

One commonly used technique to study synaptic function is fluorescence imaging, which allows for the visualization and tracking of molecules within synapses [3]. For example, the use of genetically encoded fluorescent proteins can allow for the labeling and imaging of specific types of synaptic proteins, such as neurotransmitter receptors or synaptic vesicles. By combining experimental, analytical, and computational techniques, we gain insights into the behavior of synapses and the underlying patterns that govern their activity, which can lead to a better understanding of the functioning of the brain and the mechanisms of learning and memory [6].

2. METHODOLOGY

The hippocampus is a crucial region in the brain which plays a pivotal role in learning and memory formation. Hippocampal neurons exhibit a high degree of plasticity, with changes in neuronal activity and connectivity occurring in response to environmental stimuli and experience [7][8]. To better understand the behavior of synaptic vesicle release and reuptake (also known as the synaptic cycle), we used hippocampal neurons expressing the reporter VGLUT1-phluorin (VGLUT1-pH) in which the synaptic vesicle protein VGLUT1 is fused to the pH-sensitive fluorophore [9]. The fluorescence intensity of this reporter is low in the acidic environment of synaptic vesicles but increases dramatically when vesicles fuse with the plasma membrane in response to action potentials. Neurons were imaged in Tyrodes media containing the fuel sources of lactate and pyruvate (1.25mM each). Electrical stimulation was applied to trigger the firing of 100 action potentials (AP) at 10 Hz, 15 seconds after the start of image acquisition. We tracked changes in fluorescence intensity of neural synapses in response to electrical stimulation and plotted the changes in signal intensity, averaging the changes across all synapses (each ROI/region of interest represents a single synapse). We employed multiple graphical and computational methods to examine how individual synapses react to stimulation and forecast the resting condition level after stimulation.

To analyze the time series data, we collected fluorescence intensity data from different regions of interest and analyzed their behavior and underlying patterns. The raw value of fluorescence intensity is related to the number of collected photons collected during the experiment and can vary from experiment to experiment, depending on several factors, such as the strength of the fluorescent signal and the sensitivity of the detection camera. It is important to take this variability into account when analyzing the time series data. Here we adjust the detected fluorescence intensity by subtracting the background fluorescence intensity [7][8]. Background fluorescence, which is sometimes referred to as noise, is any signal detected beyond what is generated by the fluorochromes being measured. It comes from a variety of sources such as instrument setup and imaging parameters (excitation light or camera noise).

In this paper, we used various statistical, data visualization, and data modeling techniques to investigate two key questions related to synaptic behavior. These questions are:

1. When does synaptic vesicle release reach its peak in response to electrical stimulation?
2. Is the resting VGLUT1-pH fluorescence after stimulation related to factors such as the resting condition level before stimulation, the time interval from stimulation to peak, and the ratio of the peak fluorescence signal to the baseline fluorescence signal (F(peak)/F0)?

To address the first question, we used statistical analysis techniques to compare the time at which each synapse reached its peak fluorescence signal. This involved calculating summary statistics such as the mean and standard deviation of the peak times, and then performing statistical tests to determine whether the peak times were significantly different [10].

To investigate the second question, we used data visualization techniques to compare the resting condition signal level before and after stimulation. This involved creating line plots to compare the fluorescence intensity at different time points and then using paired t-tests to determine whether the differences in fluorescence intensity were statistically significant.

Finally, to investigate the third question, we used linear regression model to identify factors that were related to the resting condition signal level after stimulation. This involves developing a statistical model that included variables such as the resting condition level before stimulation, the time interval from stimulation to peak, and the F(peak)/F0 ratio.

3. DATA DESCRIPTION

For each stimulation, synaptic vesicle release indicated by the fluorescence intensity of VGLUT1-pH follows a pattern where the intensity increases from the resting condition level during stimulation, reaches a peak, and then drops back to the post-experiment resting level. To illustrate this pattern, we have plotted the fluorescence intensity changes of different regions of interest in a representative experiment. The data has been collected in Dr. Ashrafi's lab (https://cellbiology.wustl.edu/people/ashrafi/).

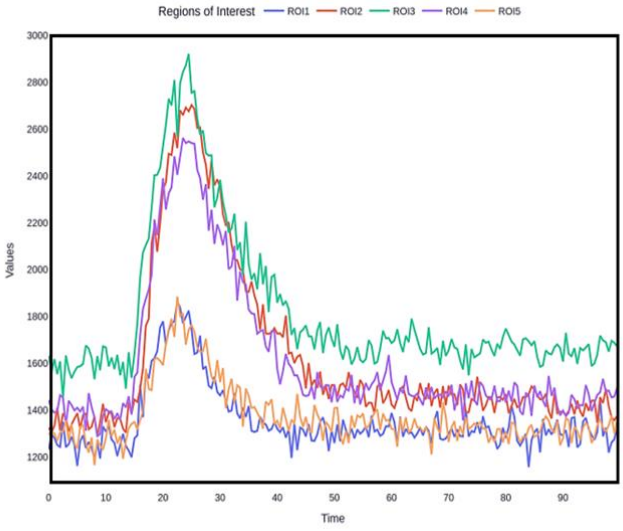

Figure 1: Time Series of ROIs

From figure 1, we observe changes in the fluorescence intensity of 5 different regions of interest (ROIs) corresponding to individual synapses in response to electrical stimulation. As electrical stimulation is applied, there is a sharp increase in fluorescence intensity in all the ROIs, indicating the release of synaptic vesicles. The time taken for the fluorescence intensity to reach its peak value also varies across different ROIs, with some ROIs reaching their peak value earlier than others. This increase in fluorescence intensity reaches a peak value and then gradually decreases back to the resting level over time.

In figure 2, we plot the time series of fluorescence intensity divided by the mean intensity level of the first 29 frames.

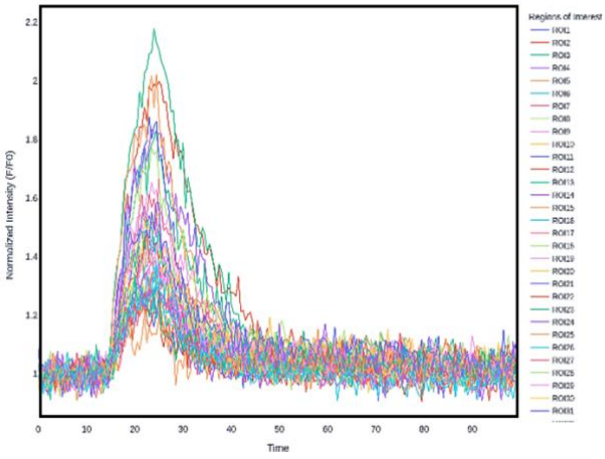

Figure 2: Normalized Fluorescence Intensity F/F0, F0: Mean Level Before Stimulation

This normalization method helps to account for any baseline differences in the fluorescence intensity levels before and after stimulation. A F/F0 value greater than 1 indicates that the fluorescence intensity at that time point is higher than the mean intensity level before stimulation. This can be interpreted as the net release of synaptic vesicle in response to stimulation.

After analyzing the time series data of fluorescence intensity changes in response to electrical stimulation for 198 sets of data. Based on table 1 the mean value of Fmax/F0 is 3.374, which means that on average, the peak value fluorescence intensity value is 3.374 times the resting level. Additionally, we found that in 75% of the experiments the ratio is lower than 3.863.

4. STATISTICAL ANALYSIS

Figure 3 provides information about the time interval required for each ROI to reach its peak fluorescence intensity from the time of electrical stimulation. To generate this figure, we used data from 198 different stimulation events, each applied to different ROIs. Based on table 1 the time interval varies from 2.5s to 15s, the average is 8.9369s, and standard deviation is 1.773s. This suggests that the time required for synapses to reach peak VGLUT1-pH level varies across different synapses. However, most of the ROIs (75%) reach their peak fluorescence intensity within 10s of electrical stimulation, indicating a relatively consistent response to stimulation across different regions of interest which is consistent with the duration of stimulation. The maximum time required for any ROI to reach peak fluorescence intensity is 15s.

5. FLUORESCENCE INTENSITY LEVEL BEFORE AND AFTER STIMULATION

For each experiment, the background intensity is different. The Florence intensity detected by the camera came both from the brain and from background. We take photos to record the background intensity data for each experiment both without and with brain. Then we used the data (with brain)- background intensity to represent the intensity change in brain to track brain activities.

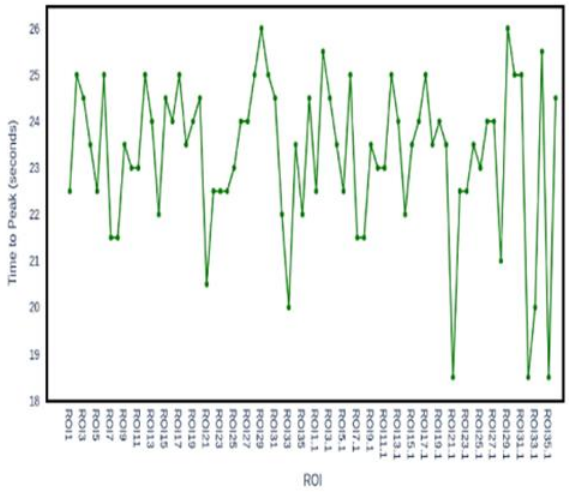

Figure 3: Time to Reach Peak at the Time of Stimulation for Each ROI

Table 1: Fmax/F0 Statistics

count	198.000000
mean	3.374195
std	1.726841
min	1.299884
25%	2.249725
50%	2.895117
75%	3.863622
max	11.718821

To determine if there is a significant difference between the resting fluorescence intensity level before and after stimulation, we employed a statistical hypothesis testing process. We used a paired statistical test, which is a type of test used when two related samples are taken, in this case, the fluorescence intensity levels before and after stimulation [11]. The null hypothesis was that the mean fluorescence intensity levels before and after stimulation were the same, while the alternative hypothesis was that they were significantly different. The significance level, or alpha value, was set to 0.05.

After conducting the statistical test, the result indicated that the null hypothesis was rejected, and the means of fluorescence intensity levels before and after stimulation were significantly different (Table 2).

Table 2: Peak Time Statistics

count	198.000000
mean	8.936869
std	1.772550
min	2.500000
25%	8.000000
50%	9.000000
75%	10.000000
max	15.000000

This implies that the stimulation had a significant effect on the fluorescence intensity level, and the difference was not due to chance. Figure 4 demonstrated that the sample is not significantly different from the population. The peak times were not significantly different.

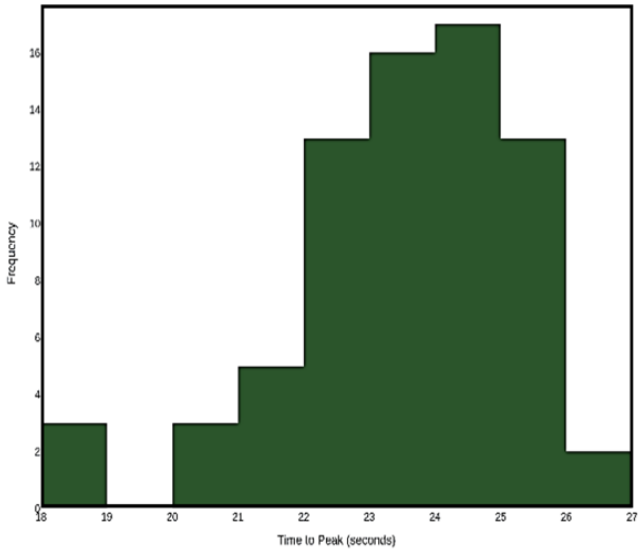

Figure 4: Histogram of Time to Reach Peak at the Time of Stimulation

Figure 5 provides a comparison of the fluorescence intensity of 36 different regions of interest (ROIs) in response to one electrical stimulation. In the plot, we can see that for most of the ROIs, the average resting condition intensity after stimulation is higher than the average resting condition intensity before stimulation. To further investigate this effect, we analyzed 198 sets of data from different stimulation. In 122 of the data sets, the average resting condition intensity after stimulation was higher than the average resting condition intensity before stimulation, while in 78 of the data sets, the average resting condition intensity after stimulation was lower. This suggests that the electrical stimulation has a positive effect on the resting VGLUT1-pH signal intensity.

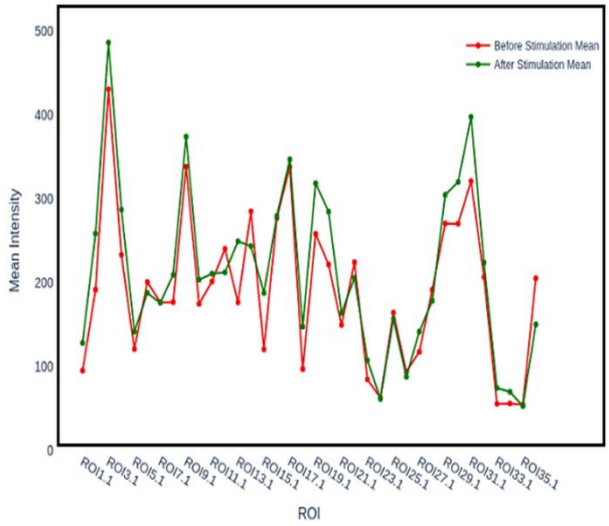

Figure 5:Mean Fluorescence Intensity Level of ROIs Before andAfter Stimulation

6. FORECASTING RESTING INTENSITY LEVEL AFTER STIMULATION

The feature df_resting_before represents a subset of the data where the time is less than 15 seconds (a threshold value representing the starting time of the stimulation). This subset focuses on the data before the stimulation period, capturing the resting state of the synapses.

On the other hand, df_resting_after is the subset of the data where the time is greater than 55 seconds. This subset captures the data after the stimulation period, representing the synapses returning to their resting state. The threshold of 55 is used to delineate the end of the stimulation and the start of the post-stimulation period.

The feature time_to_peak is created by calculating the time it takes for each synapse to reach its peak response. This is done by subtracting 15 sec from the time when the maximum value of a particular ROI occurs. The resulting values in time_to_peak indicate the time it takes for each synapse to reach its peak response relative to the start of the stimulation.

Time to peak=Time where max ROI occurs -15 sec

The feature peak represents the maximum value attained by each synapse. It provides a measure of the peak response magnitude of the synapses during the stimulation period.

Lastly, F_peak/F0 is the ratio of the maximum response (peak) to the average response during the resting period (df_resting_before.mean()).

This ratio is calculated as peak/ df_resting_before.mean(), gives an indication of how VGLUT1-pH signal changes relative to their resting state in response to the stimulation .

$$\frac{F_{peak}}{FO} = \text{peak/avg (df\_resting\_before)}$$

7. RESULTS

We visualized internal correlations between different features with a scatter plot in figure 6. The matrix shows that there is a strong positive correlation between the resting intensity level before stimulation and the resting intensity level after stimulation (0.80), indicating that the higher the resting intensity level before stimulation, the higher the resting intensity level after stimulation.

On the other hand, from table 3, there is a negative correlation between the resting intensity level after stimulation and the F_peak/F0 value (-0.36), indicating that the higher the resting intensity level after stimulation, the lower the F_peak/F0 value. Additionally, there is a negative correlation between the resting intensity level before stimulation and the F_peak/F0 value (-0.52), indicating that the higher the resting intensity level before stimulation, the lower the normalized response of synapses to electrical stimulation (F_peak/F0 value). There is no significant correlation between the time to reach peak from stimulation and other features [13].

Similarly, there is a positive correlation between the resting intensity level before stimulation and the peak value (0.45), indicating that the higher the resting intensity level before stimulation, the higher the peak value [12].

After dividing the entire dataset into two parts, we trained our model on the training dataset, which consisted of 80% of the full dataset, and tested it on the remaining 20% testing dataset.

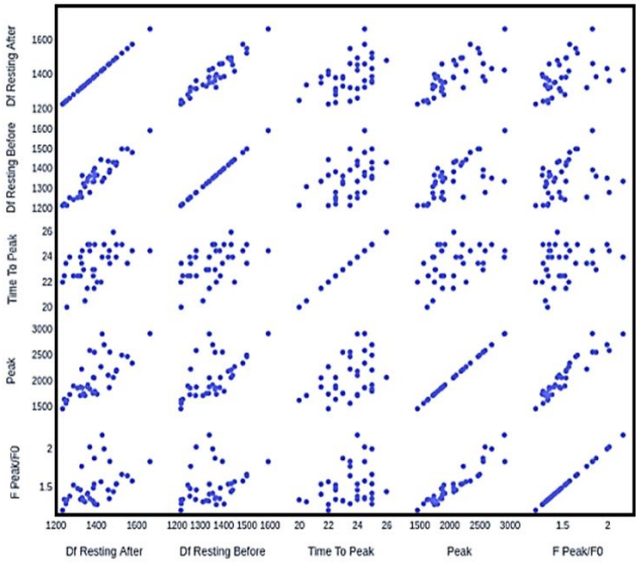

Figure 6: Scatter Plot of Key Features

Table 3: Correlation Matrix of Some Features

Correlation	df_resting_before	df_resting_after	time_to_peak	peak	F_peak/F0
df_resting_before	1	0.800965	-0.041314	0.453808	-0.524972
df_resting_after	0.800965	1	0.022777	0.442641	-0.361369
time_to_peak	-0.041314	0.022777	1	-0.05114	-0.072951
peak	0.453808	0.442641	-0.051135	1	0.392178
F_peak/F0	-0.524972	-0.361369	-0.072951	0.392178	1

Using df_resting_before, time_to_peak, and peak value as features, and df_resting_after as the target variable, we obtained regression coefficients of [0.81531833, 4.87539099, 0.03804626].

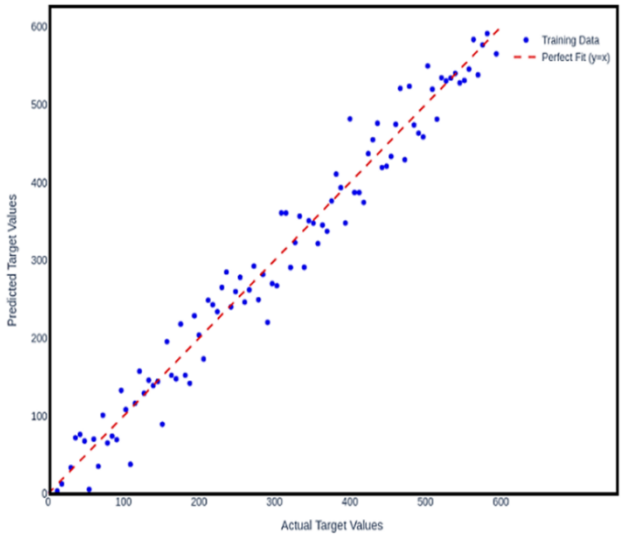

Figure 7: Enhanced Actual vs Predicted Training Set

According to figure 7, we can observe that model performance is great on the training set as most of the data points are along a straight line.

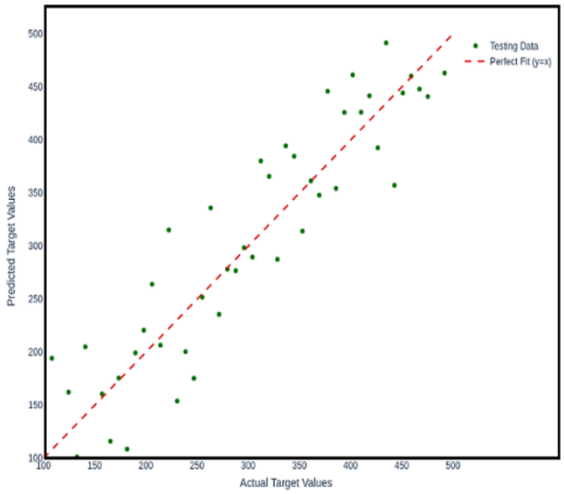

Figure 8: Enhanced Actual vs Predicted Testing Set

From figure 8, we can see that our model performed great on test set as well, with only a small amount of data points away from the straight line. To determine the significance of these features, we conducted statistical tests. Specifically, we used the F-test to determine if the overall model was significant and the t-test to determine the significance of each individual feature. Our results indicated that the overall model was significant and that all three features were significant predictors ($p < 0.05$).

Additionally, we evaluated the performance of our model by calculating the Mean Square Error (MSE) for both the training and testing datasets. The square root of the MSE for the training

dataset was 77.23639413043944, and for the testing dataset was 47.736379819011795. These results indicate that the model fits well with the training dataset and has good predictability for new data.

The regression model's coefficients for the data are presented in table 4.

8. DISCUSSION

There is a large restriction on conservation of synaptic protein centers and frames [15]. In this project, we used line plots to compare the fluorescence intensity at different time points and then using paired t-tests to determine whether the differences in fluorescence intensity were statistically significant. It helps to detect highly sensitive fluorescence and their action after stimulation.

Ther other challenge is the quality of quantitative fluorescence imaging which is affected by fluorescence signal, camera noise, and background light. In this project, Neurons were imaged in Tyrodes media containing the fuel sources of lactate and pyruvate (1.25mM each) to reduce all these problems.

9. CONCLUSIONS

In this study, we investigated synaptic vesicle release dynamics in hippocampal neurons through analysis of the fluorescence intensity of VGLUT1-pH. We used statistical analysis, data visualization, and data modeling techniques to explore three key questions related to synaptic vesicle release. Our findings indicate that the time required for synapses to reach peak activity varies slightly across different synapses, and most synapses (75%) reaching their peak fluorescence intensity within 10s, consistent with the duration of electrical stimulation.

Furthermore, we observed a positive correlation between the resting intensity level after stimulation and resting intensity level before stimulation, peak value, and time interval to peak. By applying a linear regression model, we found that the coefficient of correlation between resting_after_level and resting_before_level, time interval to peak, F_peak/F0 were 0.815, 4.875, and 0.03, respectively, with a mean squared error (MSE) of 47.736 (Table 4).

Thus, we conclude that resting VGLUT1-pH intensity before stimulation positively correlates with its level after stimulation. Furthermore, synapses with high resting VGLUT1-pH intensity reach a lower normalized fluorescence peak in response to electrical stimulation. These findings provide valuable insights into synaptic vesicle release in hippocampal synapses and enable us to understand these biological processes in quantitative detail.

List of abbreviations
ROI - Region of Interest
MSE - Mean square Error

Table 4: Dependent Variables and Standard Errors (in Parentheses)

Dependent Variable: Total Score	
Independent Variables	*Coefficients*
Intercept	73.62915 (15.712961)
Acceptance Rate	-0.1945 (0.038118)
Total Starting Compensation	0.000224 (0.000081)
Private	3.092789 (1.183021)
Average SAT	-0.017536 (0.012439)
Internship Rate	0.084531 (0.041293)
Employment Rate	0.044813 (0.065846)
Alumni Ranking A	14.404618 (1.808956)
Alumni Ranking B	10.002913 (1.614091)
Alumni Ranking C	8.042108 (1.552169)

DECLARATIONS

Ethical approval and Consent to Participate

Authors consent to participate in this project, and we know that: the research may not have direct benefit to us. Our participation is entirely volunteer. There is a right to withdraw from the project at any time without any consequences.

Consent for Publication

We give our consent for the publication of exclusive details, that could be included figures and tables and details within the manuscript to be published in Journal of Computational Neuroscience.

Competing Interests

All authors declare that they have no conflicts of interest.

Authors' contributions

B. Rahmani, Administration, Supervision of AI part, Validation, Review the paper.
H. Yiyuan, Analysis, Coding, Data Preprocessing, Visualization, Wrote the original paper.
M. Aqeel, Analysis, Coding, Visualization, Re-Wrote the original paper with more details.
P. Norouzzadeh, Validation, Methodology, Supervision of ML part, Review the paper.
A. Maazallahi, Visualization, Methodology, Review the paper.
E. Snir, Validation, Methodology, Supervision of Analysis Part, Review the paper.
G. Ashrafi, Administration, Supervision of Health and Medical Part, Idea, Review the paper.

Funding

No Funding for this project.

Availability of Data And Materials

We can upload data but should not be released by journal. Because they are calculated in Dr. Ashrafi's lab.

Acknowledgments

Not applicable

REFERENCES

[1] Yuste, R. (2015). From the neuron doctrine to neural networks. Nature Reviews Neuroscience, 16(8), 487-497. doi: 10.1038/nrn3963.

[2] Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. Perspectives on Psychological Science, 7(6), 543-554. doi: 10.1177/1745691612459060

[3] Kerren, A., Ertl, T., & Rheingans, P. (2018). Visualization and visual analysis of multimodal brain data. IEEE Transactions on Visualization and Computer Graphics, 24(1), 343-357.

[4] Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn human connectome project: an overview. NeuroImage, 80, 62-79.

[5] Broussard, G. J., Liang, Y., Fridman, M., & Uhlrich, D. J. (2018). Quantitative synaptic maps and circuits underlying critical periods of cortical plasticity. Current Opinion in Neurobiology, 54, 103-111.

[6] Buzsáki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. Hippocampus, 25(10), 1073-1188. doi: 10.1002/hipo.22488.

[7] Ramirez, S., Liu, X., Lin, P. A., Suh, J., Pignatelli, M., Redondo, R. L., ... & Tonegawa, S. (2013). Creating a false memory in the hippocampus. Science, 341(6144), 387-391.

[8] Turrigiano, G. G., & Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. Nature Reviews Neuroscience, 5(2), 97-107.

[9] Sankaranarayanan, S., De Angelis, D., Rothman, J.E., and Ryan, T.A. (2000). The use of pHluorins for optical measurements of presynaptic activity. Biophys J 79, 2199-2208

[10] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis (7th ed.). Upper Saddle River, NJ: Pearson Education.

[11] Altman, D. G., & Bland, J. M. (1994). Statistics notes: The paired t test. BMJ, 308(6933), 149-151. doi: 10.1136/bmj.308.6933.149.

[12] Jaccard, J., & Wan, C. K. (1996). LISREL approaches interaction effects in multiple regression. Thousand Oaks, CA: Sage Publications.

[13] Spiegel, M. R. (2010). Schaum's outline of theory and problems of probability and statistics (4th ed.). New York, NY: McGraw-Hill.

[14] Schwarz, L. A., Miyamichi, K., Gao, X. J., Beier, K. T., Weissbourd, B., DeLoach, K. E., ... & Luo, L. (2015). Viral-genetic tracing of the input-output organization of a central noradrenaline circuit. Nature, 524(7563), 88-92.