Benchmarking Deep Learning Architectures for Predicting Visual Stimuli Given Single Neuron Spike Patterns

Raymond Carpenter (rcarpenter@unc.edu)

Department of Statistics & Operations Research, University of North Carolina at Chapel Hill, 318 E Cameron Ave
Chapel Hill, NC 27599, USA

Cole Vita (cole.vita@unc.edu)

Department of Statistics & Operations Research, University of North Carolina at Chapel Hill, 318 E Cameron Ave
Chapel Hill, NC 27599, USA

Vladas Pipiras (pipiras@email.unc.edu)

Department of Statistics & Operations Research, University of North Carolina at Chapel Hill, 318 E Cameron Ave
Chapel Hill, NC 27599, USA

Abstract:

Drawing insights from neuronal processes is integral for understanding the neural mechanisms underlying cognitive processes, providing a higher definition recording for brain-computer interfaces, and helping develop advanced neurorehabilitation strategies. Our study sought to survey and identify machine-learning models and deep-learning architectures capable of predicting visual stimuli based on the spike patterns of single neurons. We worked with Neuropixel data from the Allen Brain Observatory [1,2] consisting of the firing rates from single neurons, also called spike trains, from several male mice's visual cortex, thalamus, and hippocampus. Each recording involved around 2,000 separate units. The mice were shown 118 different natural images of predators, foliage, and other scenes from their natural habitat at random in repetition and for 250 ms each. The firing rates of the separate units were then used as predictors for the shown images. A Long Short-Term Memory (LSTM) network provided the highest accuracy, with up to 96.6% accuracy. Other architectures, such as Transformer networks and Graph Attention Networks, had prediction accuracies of over 90%.

Keywords: Function Prediction; Visual Processing; Deep Learning; Spike Train Data

Introduction

Drawing insights from neuronal processes is integral for understanding the neural mechanisms underlying cognitive processes, providing a higher definition recording for brain-computer interfaces, and helping develop advanced neurorehabilitation strategies. Our study sought to survey and identify machine-learning models and deep-learning architectures capable of predicting visual stimuli based on the spike patterns of single neurons. We worked with Neuropixel data from the Allen Brain Observatory [1,2] consisting of the firing rates from single neurons, also called spike trains, from several male mice's visual cortex, thalamus, and hippocampus. Each recording involved around 2,000 separate units. The mice were shown 118 different natural images of predators, foliage, and other scenes from their natural habitat at random in repetition and for 250 ms each. The firing rates of the separate units were then used as predictors for the shown images. (Figure 1)

Results

We assessed the prediction performance on test data for various machine and deep learning architectures built on training data. A random guess was associated with a baseline test accuracy of $1/118 = 0.85\%$. Support Vector Machine and Principal Component Regression had minimal success. A single-layer neural network (NN) on aggregate firing rates over the length of a visual stimulus resulted in a test accuracy of 93%. The test accuracy of multi-layer NNs would diminish for each layer added. To test the utility of spatial modeling, a single-layer Graph Convolutional Network (GCN) and a graph attention (GAT) network were tried with 48% and 89.8% accuracies, respectively. A Long Short-Term Memory network was tested to consider the temporal aspect of the data. The firing rates

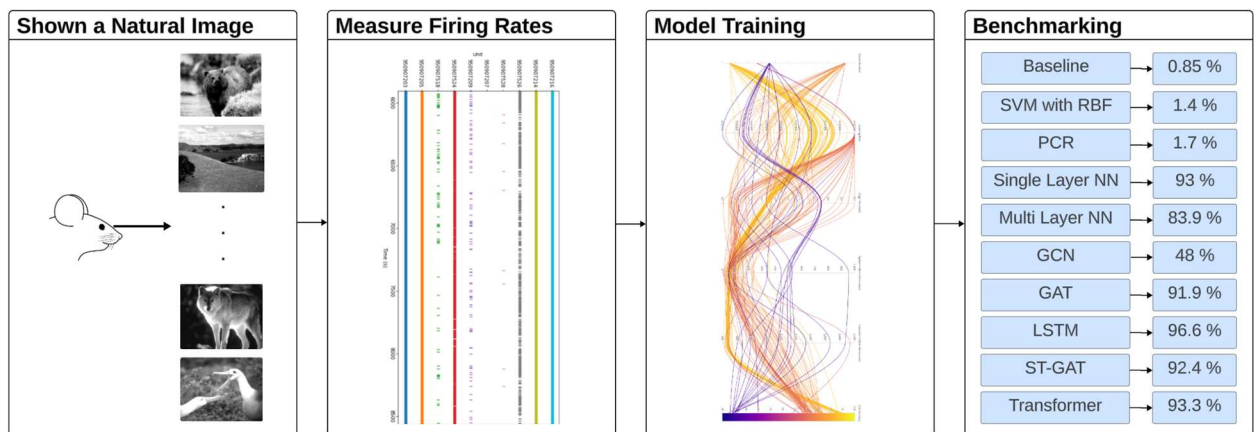| Shown a Natural Image	Measure Firing Rates	Model Training	Benchmarking	
			Baseline	0.85 %
			SVM with RBF	1.4 %
			PCR	1.7 %
			Single Layer NN	93 %
			Multi Layer NN	83.9 %
			GCN	48 %
			GAT	91.9 %
			LSTM	96.6 %
			ST-GAT	92.4 %
			Transformer	93.3 %

Figure 1: Methodology overview and model performance for predicting visual stimuli from neuronal spike trains.

during a single stimulus were broken into ten time bins for models with a temporal component. The LSTM produced the highest test accuracy at 96.6%. A transformer was also tested, which may account for the spatial aspect through its attention mechanisms and the sequential nature of the data. This had a related test accuracy of 93%.

A Spatial-Temporal Graph Attention Network (ST-GAT) was also implemented to account for both the spatial and temporal aspects. It accounts for the spatial aspects through the first layer, the GAT. The output is then passed through the LSTM to account for temporal data. The ST-GAT produced a test accuracy of 92.4%. This model allows for an adjacency matrix to be found through backpropagation, potentially representing functional connectomics between single neurons [3]. (Figure 2)
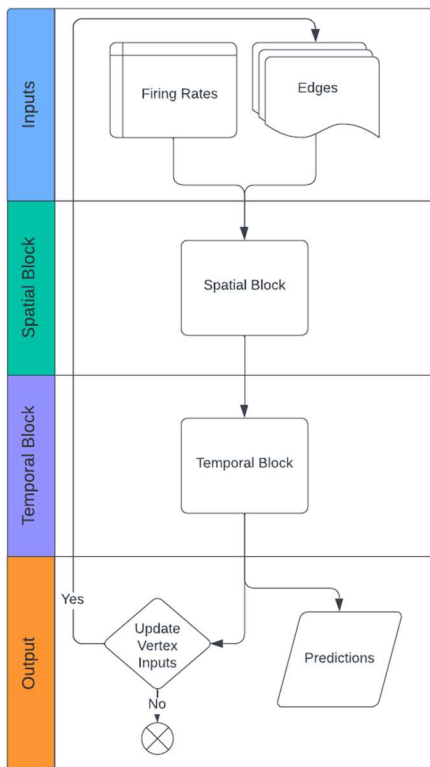

Figure 2: An architecture overview of a Spatial-Temporal Graph Attention Network (ST-GAT). Showing the ability to find graph edges during training. In this case the edges are the connection between single neurons or units.

Discussion

We have found that architectures with fewer layers, including NNs, LSTMs, and Transformers, consistently demonstrated higher test accuracies than their multi-layered counterparts, which had many more parameters and likely overfitted by capturing noise in the data. Additionally, the success of LSTM and Transformers suggests that including a temporal component allows models to handle the sequential nature of neural data, increasing prediction accuracy. Our results were consistent across several mice.

Both machine learning models, Support Vector Machines and Principal Component Regression, performed worse than all other deep learning-based models. The ability to tune multiple weights to specific neurons might account for the relative success of deep learning architectures over machine learning models. This supports their strength in modeling spike train data.

Our results provide key insights into the compatibility of different learning architectures in drawing valid conclusions about function from neural data at the micro-level. Single-layer neural Networks, Transformers, LSTMs, and Spatial-Temporal Graph Neural Networks can all predict visual stimuli with up to 96.6% accuracy. Given related neural responses, similar modeling may be extended to predict other functions in mice and humans more broadly.

The code for this project can be found at the following repository: https://github.com/RayCarpenterIII/Neuropixel-Analysis

References

Allen Brain Observatory. Neuropixels Visual Coding. https://portal.brain-map.org/circuits-behavior/visual-coding-neuropixels

Paulk AC, Kfir Y, Khanna AR et al. Large-scale recording with single neuron resolution

using Neuropixels probes in human cortex. Nature Neuroscience, 2022, 25: 252-263.

Wein S, Schuller A, Tome AM, et al. Forecasting brain activity based on models of spatiotemporal brain dynamics: A comparison of graph neural network architectures. Network Neuroscience, 2022, 6 (3): 665–701.

Ray Tune: Hyperparameter Tuning. https://docs.ray.io/