NX414 Project - Report

Myriam Hamon, David Martin-Osorio

1 Introduction

This project investigates different models for neural activity prediction using linear regression, starting from images and going to different models either data-driven or model-driven. This aims to both create the best model possible and understand why that model might be the most predictive of the neural activity recorded in the Inferior Temporal cortex (IT) of monkeys with multielectrode arrays from Majaj et al. [1].

2 Investigation of the difference in performance over models

After evaluating various approaches for predicting IT neural activity, we found that goal-driven models and especially models, retrained on image classification tasks and fine- tuned by adapting their final layers for spike prediction yielded the best results (see Table 1) , using a common machine learning technique knows as transfer learning [2].

Inputs	Regression	Score	Explained Variance	Average Correlation
Images				
Raw	Linear	-0.037	-0.033	0.052
Raw	Ridge	-1.167	-1.157	0.060
Top 1000 PCs	Linear	-0.089	-0.085	0.053
Top 1000 PCs	Ridge	-0.089	-0.085	0.054
Top 1000 PCs	Ridge CV	0.086	0.090	0.055
Task-driven models				
Resnet				
pretrained weights	Ridge CV	0.337	0.339	0.113
random weights	Ridge CV	0.029	0.033	0.089
Goal-driven models				
Shallow CNN				
4 layers	Ridge CV	0.24	0.243	0.11
6 layers	Ridge CV	0.23	0.23	0.128
Resnet				
not pretrained	Ridge CV	0.261	0.267	0.104
pretrained	Ridge CV	0.395	0.40	0.102
AlexNet				
pretrained	Ridge CV	0.39	0.39	0.12
VGG16				
pretrained	Ridge CV	0.39	0.40	0.11

Table 1: Performance of regression on validation data using different inputs. For model-based inputs, regression was performed on the top 1000 principal components (PCs) of all or selected layer activations. The reported performance corresponds to the layer yielding the best regression results.

3 Investigation of the difference in performance within the layers of a task-driven model

Layer	Score	Explained Variance	Average Correlation
Conv1	-0.005	0.00	0.045
Layer1	0.113	0.115	0.094
Layer2	0.243	0.246	0.106
Layer3	0.337	0.339	0.113
Layer4	0.286	0.191	0.099
Avgpool	0.001	0.00	0.073

Table 2: Performance of the ridge regression (cross-validated) of the activations of different layers of the ResNet-50 for predicting spikes.

In the pretrained ResNet, a task-driven model, performance varied notably across layers. The first and last layers were the least predictive of spiking activity, likely due to their representations being too low-level or overly task-specific. For example, he final layer (avgpool) shows nearly zero explained variance across all neurons (Figure 1), suggesting that its features carry little to no information relevant to IT spiking activity. These findings align with Yamins et al. [3], who showed that deeper network layers correspond to higher-level visual areas.

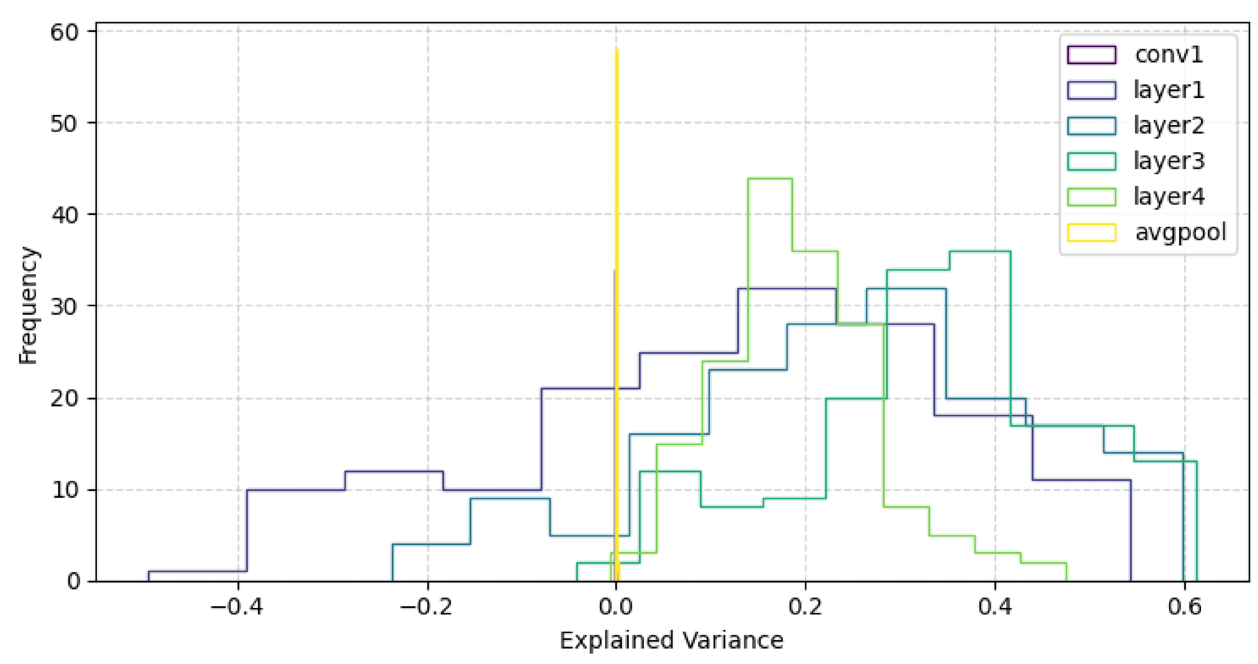

Figure 1: Distribution of the explained variance of each layer of the ResNet-50

4 Results

The ResNet model emerged as the most effective, excelling in both spike prediction accuracy and image classification performance. To enable a more comprehensive comparison, it would be beneficial to implement stricter controls on the number of parameters retrained. The retrained network can be find at `https://drive.google.com/file/d/1OIOwfslgsR0LRVzZ0BCx4YLSh5NziJJS/view?usp=sharing`

Moreover, exploring an optimization strategy to selectively retrain specific layers or parameters could offer valuable insights into the models' behavior and further enhance their performance.

Given more time, we would have liked to investigate more advanced architectures and explore additional optimization techniques, such as hyperparameter tuning and different regularization methods, to improve the current models.

References

[1] Najib J. Majaj et al. "Simple Learned Weighted Sums of Inferior Temporal Neuronal Firing Rates Accurately Predict Human Core Object Recognition Performance". In: *Journal of Neuroscience* 35.39 (2015), pp. 13402–13418. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.5181-14.2015.

[2] Sinno Jialin Pan and Qiang Yang. "A Survey on Transfer Learning". In: *IEEE Transactions on Knowledge and Data Engineering* 22.10 (2010), pp. 1345–1359. DOI: 10.1109/TKDE.2009.191.

[3] Daniel L K Yamins and James J DiCarlo. "Using goal-driven deep learning models to understand sensory cortex". In: *Nature Neuroscience* 19.3 (Feb. 2016), pp. 356–365. ISSN: 1546-1726. DOI: 10.1038/nn.4244. URL: http://dx.doi.org/10.1038/nn.4244.