NX-414 MiniProject Report

Florence Crozat, Ernesto Bocini

***Abstract*—The project consisted in trying to predict neural activity from images. It is an example of a data-driven approach to predict brain activity. The metrics used during this project to compute the goodness of the model were explained variance and correlation between real and predicted values.**

I. Introduction

In this project we use and analyse data from the paper by *Najib J. et al*[1]. The behavioral experiment explained in the paper consisted in showing to non-human primates some images while recording the neural activity with multielectrode arrays from the inferior temporal (IT) cortex. Here, the neural activity and the images are already pre-processed and we analyze try to explain the neural activity (firing rate between 70 and 170 ms) using different stimulus images per each neuron.

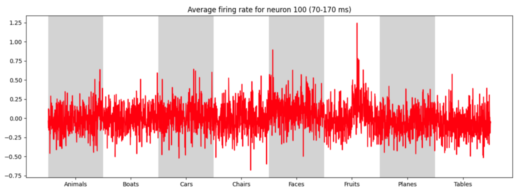

Fig. 1. Average response for neuron 100 dividend into object types

II. Ridge Regression and PCA

The first model used linear regression to predict neural activity from single pixels, which yielded mitigated results – as expected since predicting from pixels is very hard and the model is overfitting. Specifically, the model chosen was Ridge Regression. We tried to enhance this prediction using one-hot encoder for each class, but the explained variance by the model stayed negative – meaning that the model made worse predictions than random predictions.

In order to reduce overfitting, we then investigated the effect of reducing the number of components through PCA, to retain only 1000 components. That improved the explained variance, but the model would still not predict neural activity as at best we could explain was 40% for some specific neurons and the average explained variance stayed very low: only 4.6 %. We tried to enhance these results by using cross-validation to find the regulation term yielding the better prediction according to the metrics. However results were upper-bounded by the previous finding and on average were even worse than without cross-correlation. Numerical values for the metrics are described in Fig. 3.

III. Using a pretrained model

We then investigated from the pre-trained **ResNet50** model and compared predictions from some of the layers once PCA done. The PCs of the different layers were then fed into the previously created linear model to predict neural activity. It led to different prediction goodness, as shown in Fig. 2.

The results were also compared with randomly initialized weights instead of the current best weights trained for the model. Numerical results of this comparison are presented in the summary table Fig. 3.

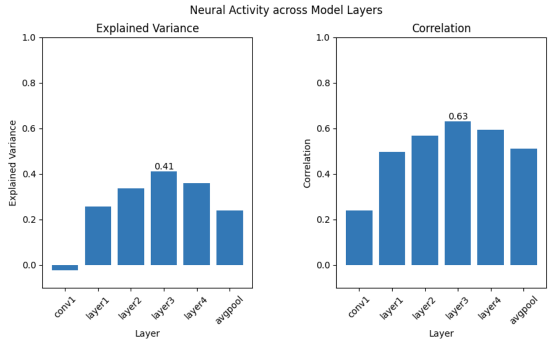

Fig. 2. Explained variance and correlation across layers for the ResNet50 pretrained model

IV. Challenge

In order to develop the most accurate prediction for neural activity from pixels, we tried different approaches.

One of them aimed to improve the results of the ShallowCNN using the Python library Optuna, which aims to find the best hyperparameters for a specific model. In our case, the hyperparameters chosen were the optimizer (chosen between Adam and SGD) as well as the learning rate. The tuning aimed at increasing the explained variance. The results improved from an explained variance of roughly 0% before using Optuna to an explained variance of 15.7% afterwards. The correlation improved from again 0% to about 40% after hyperparameters tuning.

In order to achieve an even better performance, we then decided to kepp working on the Resnet50 model defined in Part III, and to feed the complete result to a fully connected layer to predict activations. ResNet50 is a deep convolutional neural network architecture that consists of 50 layers. It uses skip connections to allow gradients to propagate more effectively during training and has achieved state-of-the-art performance on a variety of image recognition tasks. In this code, the final fully connected layer of the ResNet50 model is replaced with a new layer that is specifically designed for the regression task, where the goal is to predict the spiking activity of 168 neurons in response to a given visual stimulus. The model obtained an explained variance of 44%. Also other models such as *Cornet-Z* and *Cornet-S* (*M. Schrimpf et al* [2]) but without similar success.

V. Results Summary

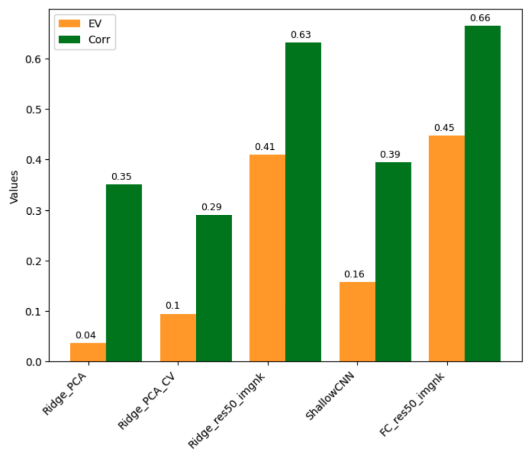

Fig. 3. Summary table of metrics evaluating the goodness of different models

REFERENCES

[1] Najib J Majaj, Ha Hong, Ethan A Solomon, and James J DiCarlo. Simple learned weighted sums of inferior temporal neuronal firing rates accurately predict human core object recognition performance. *Journal of Neuroscience*, 35(39):13402–13418, 2015.

[2] Jonas Kubilius, Martin Schrimpf, Aran Nayebi, Daniel Bear, Daniel LK Yamins, and James J DiCarlo. Cornet: Modeling the neural mechanisms of core object recognition. *BioRxiv*, page 408385, 2018.