AFRL-AFOSR-VA-TR-2021-0034

Neuromorphic models of the visual system for multichannel, spike-based encoding and processing.

Perez-Acle, Tomas
FUNDACION CIENCIA PARA LA VIDA
AV. ZANARTU 1482 NUNOA
SANTIAGO, SANTIAGO, 99999
CHL

04/22/2021
Final Technical Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* 22-04-2021	**2. REPORT TYPE** Final	**3. DATES COVERED** *(From - To)* 15 Jul 2019 - 14 Jul 2020

4. TITLE AND SUBTITLE Neuromorphic models of the visual system for multichannel, spike-based encoding and processing.	**5a. CONTRACT NUMBER** FA9550-19-1-0368
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Tomas Perez-Acle	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) FUNDACION CIENCIA PARA LA VIDA AV. ZANARTU 1482 NUNOA SANTIAGO, SANTIAGO 99999 CHL	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AF Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/AFOSR IOS
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-AFOSR-VA-TR-2021-0034

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The mammalian retina is a complex neural tissue composed by hundreds of millions of specialized neurons. Even though its architecture is simpler than that of the brain cortex, it processes a variety of visual signals by optimally filtering and extracting relevant environmental information through complex computing. Visual signals processed by the retina are then transmitted to the brain in the form of spikes through several parallel channels -the neuronal axons- forming the optic nerve, where they are further processed in a parallel and hierarchical way, at increasing levels of complexity. Recent research has shown that the number of information channels emerging from the retina, defined by the number of different Retinal Ganglion Cells types, is much larger than previously thought. While functionally characterized, they are still not fully understood, and defining why so many distinct channels are present is still an open question. Building upon a set of artificial neural networks (ANN) models accounting for the encoding process of nearly 40 types of Retinal Ganglion Cells (RGCs), we studied the response of each ANN model -we termed Neuromorphic Neural Circuits (NNCs)- to different images. In doing so, we characterized how much and what information is
shared between these NNCs representing the variety of RGCs. We found that despite our NNCs being trained with artificial images, it is capable of replicating some of the expected behavior of the retina when stimulated with natural images. Even though there is considerable evidence supporting the notion that the mammalian visual system is evolutively tuned to optimally encode natural images, our work is the first one to perform a direct comparison between artificial and natural images using the complete set of retinal outputs.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** DANIEL MONTES
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)* 000-0000
U	U	U	SAR	9	

Neuromorphic models of the visual system for multichannel, spike-based encoding and processing (NeMoV).

Grant number FA9550-19-1-0368

Ravello, César R.; Perez-Acle, Tomas O.
Computational Biology Laboratory - Fundación Ciencia & Vida

Summary

The mammalian retina is a complex neural tissue composed by hundreds of millions of specialized neurons. Even though its architecture is simpler than that of the brain cortex, it processes a variety of visual signals by optimally filtering and extracting relevant environmental information through complex computing. Visual signals processed by the retina are then transmitted to the brain in the form of spikes through several parallel channels -the neuronal axons- forming the optic nerve, where they are further processed in a parallel and hierarchical way, at increasing levels of complexity.

Recent research has shown that the number of information channels emerging from the retina, defined by the number of different Retinal Ganglion Cells types, is much larger than previously thought. While functionally characterized, they are still not fully understood, and defining why so many distinct channels are present is still an open question.

Building upon a set of artificial neural networks (ANN) models accounting for the encoding process of nearly 40 types of Retinal Ganglion Cells (RGCs), we studied the response of each ANN model -we termed Neuromorphic Neural Circuits (NNCs)- to different images. In doing so, we characterized how much and what information is shared between these NNCs representing the variety of RGCs. We found that despite our NNCs being trained with artificial images, it is capable of replicating some of the expected behavior of the retina when stimulated with natural images. Even though there is considerable evidence supporting the notion that the mammalian visual system is evolutively tuned to optimally encode natural images, our work is the first one to perform a direct comparison between artificial and natural images using the complete set of retinal outputs.

Future projections of this work will move forward on the functional characterization of these retinal circuits and, in turn, advancing on the understanding of the relationship between retinal encoding and cortical responses of the primary visual cortex and, subsequently, with higher visual cortical areas as well.

Introduction

Understanding how populations of neurons process sensory information to generate appropriate behavioral response is one of the main goals of system neuroscience. Recent technical advancements allowed the simultaneous recording of the activity in large populations of neurons, either by imaging fluorescent probes (Euler et al. 2009) or by directly measuring the electrical activity with multielectrode arrays (Litke et al. 2004). While these techniques require expensive equipment and the capabilities to manage large amounts of data, recent trends and initiatives have motivated many scientists to make their data publicly available for the whole scientific community to use.

An essential tool in trying to understand neural behavior is the use of mathematical/computational models to explain the input/output relationship in a single cell or in a group of similar cells. However, with the increase in data availability, an emerging challenge is the development of efficient algorithms to fit models of large collections of different neuron types. In this regard, deep learning techniques have already been proven to be an effective approach in replicating the responses of the visual system at different stages (McIntosh et al. 2016, Batty et al. 2016). Despite their applicability, they have been limited to account only to a reduced number of different types of cells, particularly when it comes to model the neuronal encoding of the mammalian retina.

We proposed the development of a set of Neuromorphic Neural Circuits (NNCs) that replicate the response of each type of Retinal Ganglion Cells (RGCs) to study their response properties. These models were built using artificial neural networks tools, specifically, Long Short-Term Memory (LSTM) units.

The main objective of this project was to investigate the role of the multiple types of retinal encoders (represented by the response of the different types of ganglion cells), in terms of what aspect or property of the image is represented in each channel. This work plan was divided into three specific objectives:

Objective 1: characterize the response of a collection of neuromorphic models of the retina to different stimuli.

Objective 2: determine the level of redundancy or correlation between channels from the point of view of information theory.

Objective 3: estimate the information encoded by the whole set of retinal encoders.

We will present results regarding the first two objectives. Result for Objective 3 are currently under development, mostly due to delays caused by the COVID19 pandemic, where a considerable part of our team was temporarily reassigned to tasks related to monitoring and forecasting the spread of the disease on our country, and designing containment strategies, work that was directly reported to the Minister of Science, and that will also appear in two scientific papers, one under review and available as preprint, and one in preparation.

Thanks to a new funding coming from our research institute, we were able to hire new people for these tasks, thus most of the team was able to return to their previous lines of research, including the one reported here. We expect that the results pertaining to Objective 3 will be completed very soon, which will in turn allow us to publish our results.

Methods

We evaluated the response to different types of stimulus of 39 different types of modeled Retinal Ganglion Cells (RGCs), which would constitute most of the output of the retina. The 39 models were built using the same artificial neural network architecture, each trained with a collection of pairs of stimulus/response of neurons of each type from the Retinal Functomics database (Baden et al. 2016; Franke et al. 2017). The architecture is a relatively small network consisting of a fully connected layer, followed by two Long Short-Term Memory layers, and finally a Poisson process to transform firing rates into spike trains, implemented and trained using the PyTorch framework.

Stimuli used to test the encoding properties of the NNCs can be classified into two main categories, simple synthetic stimuli and complex natural stimuli. The simple stimuli are the typical for vision research, including drifting sinusoidal gratings, gaussian noise and chirp, which is a recently introduced pattern consisting of two consecutive oscillating sequences, one modulated in amplitude and one modulated in temporal frequency, preceded by a square wave, and that has been widely adopted for classification of cells according to their response profile.

The second set of stimuli corresponds to photographs and sequences of images taken from videos recorded in natural environments.

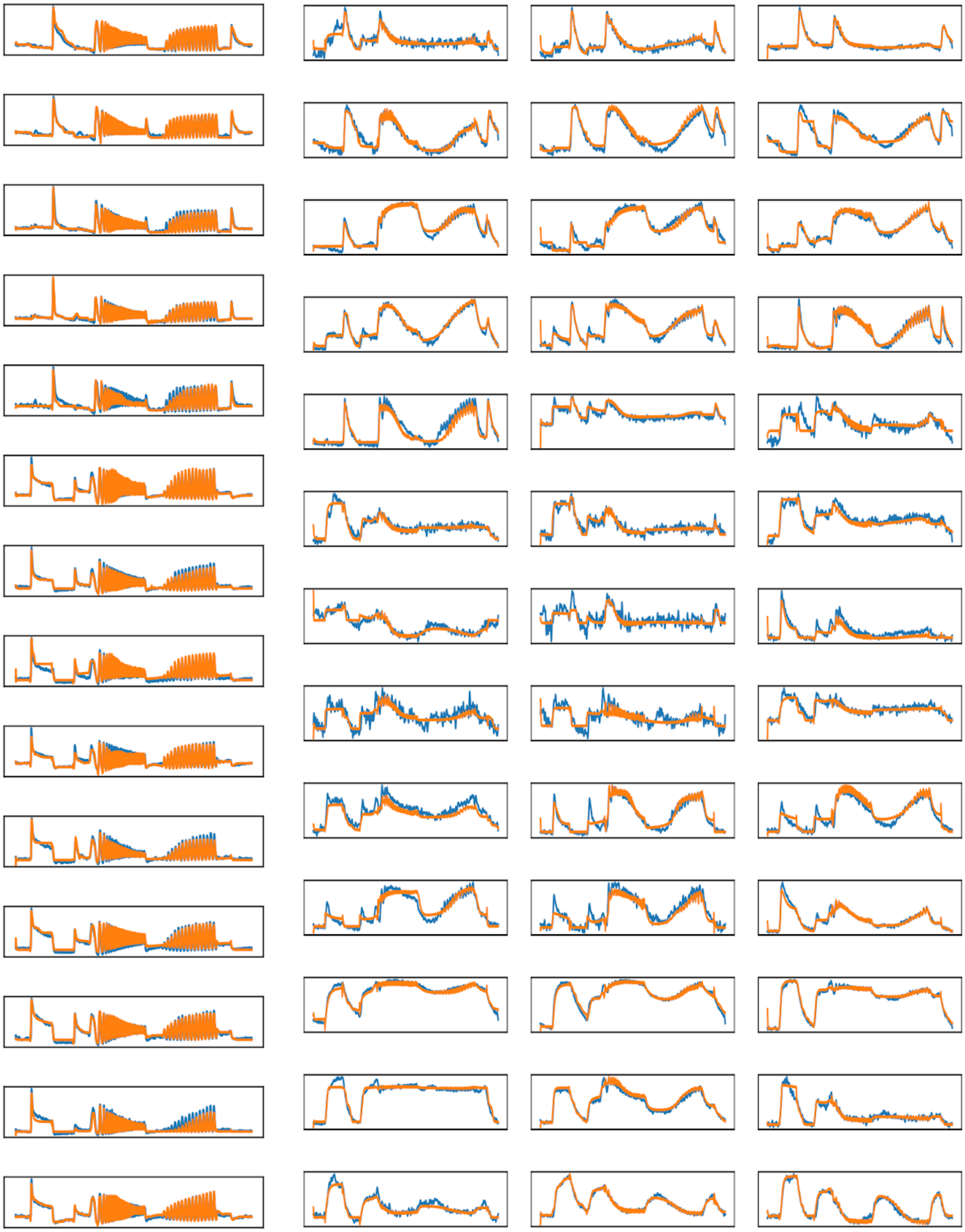

Figure 1: The Neuromorphic Neural Circuits (NNCs) replicate the response of each type of retinal neuron. The first column shows the response of the bipolar cells, while the rest show the response of the RGCs. The blue line is the neural response to the chirp stimulus, while the orange line is the response of

the NNC.

Results and Discussion

Figure 1 shows the response of each type of NNC model to the chirp stimulus. While it is a relatively simple stimulus, by sweeping over a range of frequencies and intensities, it allows to differentially activate neurons with different receptive field properties. More interestingly, it allows for a properly trained model to capture the neural response to the different properties of the stimuli, resulting in models capable of rich responses to other, non-trained stimuli, as will be seen below.

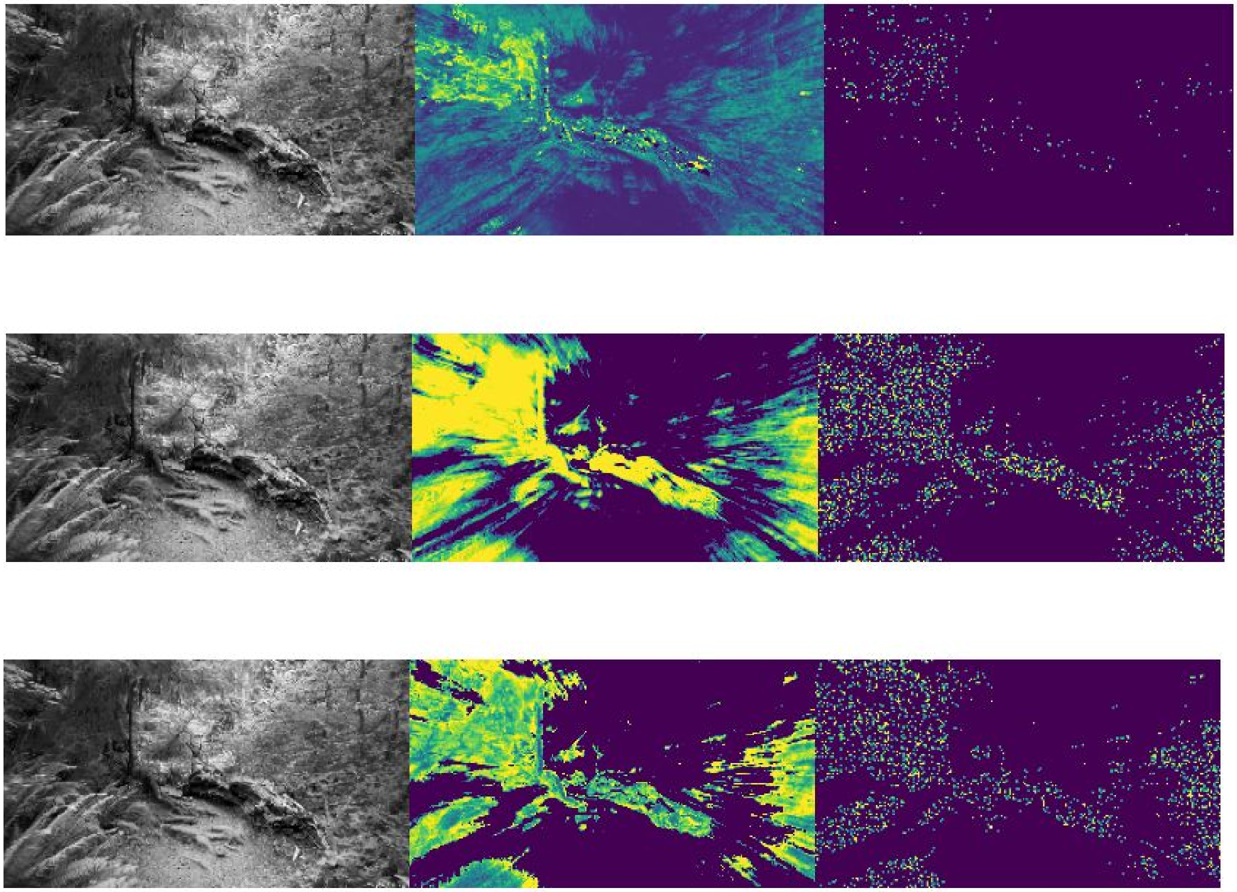

Figure 2: Response of the model to the natural images. The column shows the input image, the middle column shows the output of the different neurons (one neuron per pixel) and the right column shows the response in spikes. Each row corresponds to a different type of neuron.

Figure 2 shows some examples of the response to the natural image sequence, where it can be seen that different neurons respond to different parts of the scene. These differential responses are essential for the coding properties of the visual system. As has been proposed in the literature and backed by considerable evidence, the visual system in animals has evolved to efficiently encode the properties of the natural scenes that are relevant to the animal (Butts & Goldman, 2006). One of the implications of this efficient code is that redundancy in the information transmitted is largely reduced. This reduction of redundancy manifests in the information transmitted by neighboring neurons of the same type, and in the information transmitted by neurons of different types in the same location. To isolate those effects, we worked with a simplified model where there is no spatial information in the stimulus nor in the neural circuit architecture, focusing on the second type of redundancy reduction.

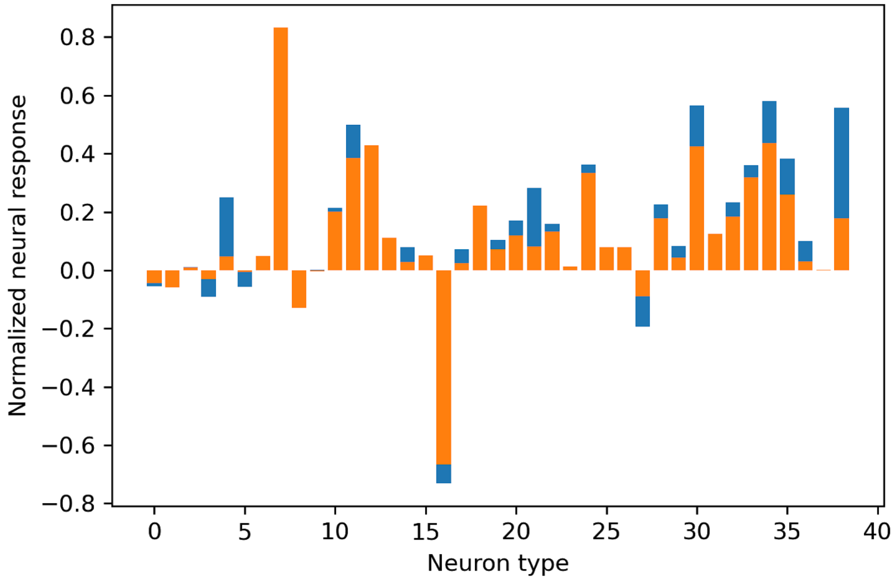

Figure 3: Response of each cell type. The bar chart shows the average response to a sequence of stimulus. The blue bar is the response to the natural stimulus, the orange bar is the response to the Gaussian noise. The differences in the response are not significant (p=0.052, Wilcoxon test).

To focus on the coding properties of our models, we compared the response of each RGC type NNC to different sequences of natural images with the response to gaussian noise. While both stimuli elicited equivalent responses in terms of firing rate (Figure 3, p=0.052, Wilcoxon test), there were differences in the response patterns when comparing different cell types. Following the idea that if different neurons generate a similar response pattern to an stimuli, then the redundancy of the information encoded by each of those neurons will be high (Barlow, 2001). To test that idea, we analyzed the Pearson correlation between the responses of all the neuron types, and found that for the Gaussian noise stimulus (for which the visual system is not tuned) the pattern of correlations between neuron types is significantly higher than for natural scenes (see Figure 4).

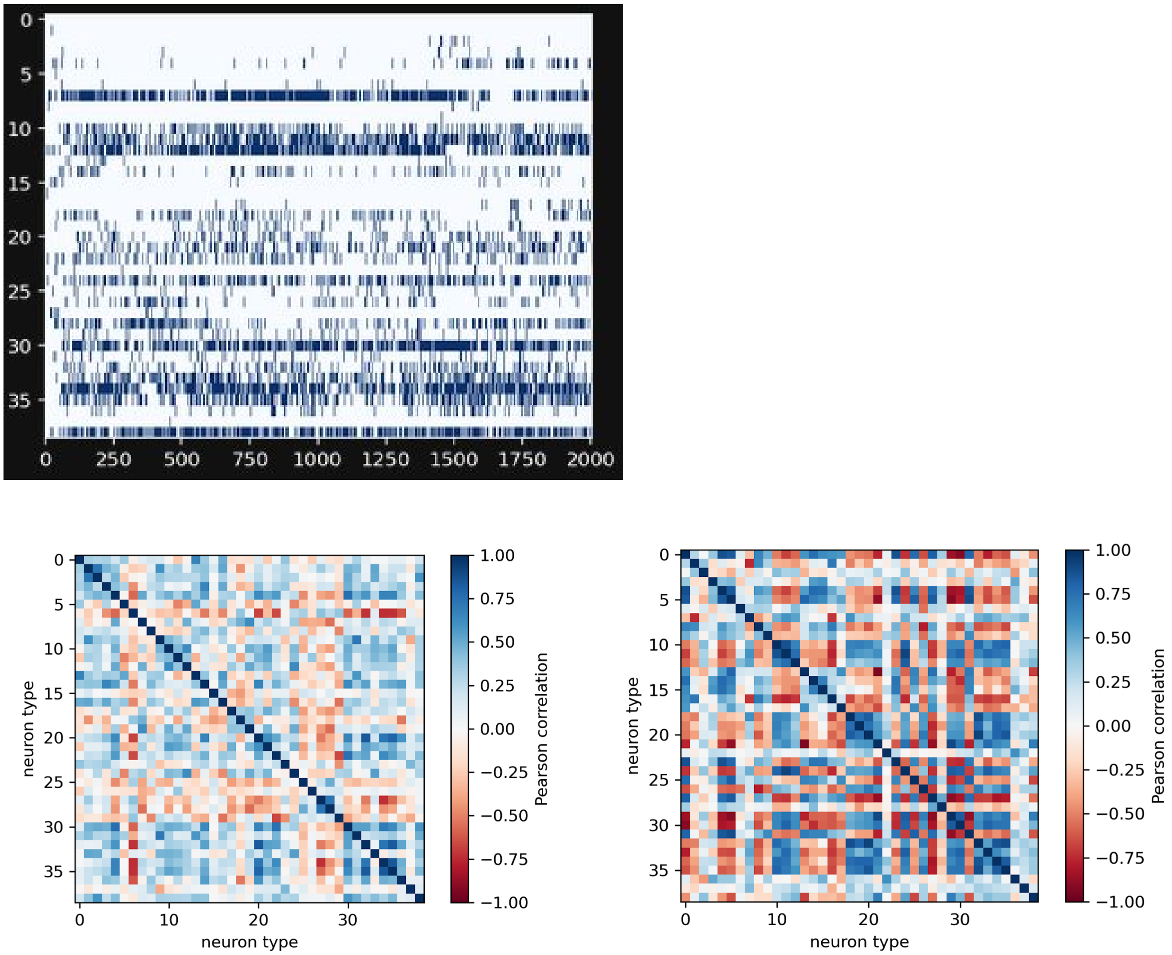

Figure 4: Differences in the response of each type of neuron. The upper panel shows the spike trains elicited by the same stimulus on each cell type. Each row represents the response of a neuron along time, the lines denote when the cell fired. As can be seen, the same stimulus can generate very different response patterns. The lower panels show the correlation matrix between each type of neuron; the panel on the left shows the correlation in response to the natural stimulus, the panel on the right the response to the Gaussian stimulus. As can be seen, the noise stimulus results in higher correlation between neural responses.

To quantify the redundancy in the information transmitted by the different neuron types, we calculated the Maximal Information Coefficient (MIC) between the response of neurons to the same set of stimuli. As can be seen in Figure 5, the response to Gaussian noise has significantly higher MICs than the response to natural images ($p = 0.003$, Wilcoxon test), thus the redundancy in the information encoded in the neural responses is lower for the natural stimuli, for which the visual system has evolved to efficiently encode.

This is a novel result that has not been previously reported. Current work is focused in analysing in more detail how the stimulus information is distributed across these different encoding channels.

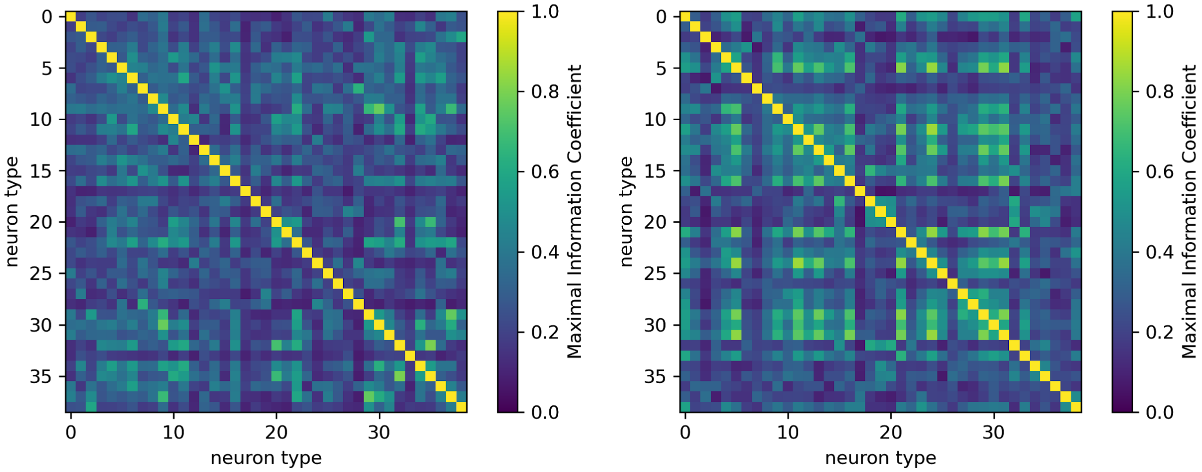

Figure 5: Maximal Information Coefficient between neural responses. The panel on the left shows the MIC matrix for the natural stimulus, the panel on the right shows the MIC matrix for the Gaussian noise. The Gaussian noise response has higher values of MIC, thus more redundancy in the encoding, while the natural images elicit responses with lower MIC, meaning that the encoding of those images is more efficient.

Conclusions

Thanks to this project, we proved that a somewhat simple temporal stimulus pattern (the "chirp") is enough to train models to replicate the responses of almost all the RGC types in the retina. This simplified model was able to reproduce behavior that was not part of the training, namely the tuning for natural images to produce efficient encoding. For the first time, we found evidence that the complete set of retinal encoders, despite its large number and apparent similarity in the response to simple stimuli, elicit distinguishable responses when stimulated with the signals that it is tuned to.

The current understanding of the subject is that models tend to become better predictors when trained on natural images than when trained with simple artificial stimuli, however, we have found that a simple yet rich stimulus is able to train models to replicate neural behavior that until now was thought to be exclusive to natural images. We have shown that the temporal tuning of the neurons is enough to partially explain the response properties to natural images. We expect that these results will be published promptly. Future research will be focused on studying the contribution of spatial tuning to the response to natural images, and then pursue a spatiotemporal framework to explain the tuning properties of retinal neurons.

Main outcomes of the project

This project helped to consolidate the line of research at the laboratory, allowing us to apply and receive additional funding from other agencies to continue the work for the immediate future. This research also attracted the interest of many potential students, so for this year we recruited three Ph.D. students that will do their thesis on this subject, and we expect to incorporate more students and possibly a postdoctoral researcher in the coming years.

References

Baden, T., Berens, P., Franke, K., Rosón, M., Bethge, M., Euler, T. (2016)**. The functional diversity of retinal ganglion cells in the mouse** https://dx.doi.org/10.1038/nature16468

Barlow, H. (2001)**. Redundancy reduction revisited.** Network (Bristol, England) 12(3), 241 - 253.

Batty, E., Merel, J., Brackbill, N., Heitman, A., Sher, A., Litke, A., Chichilnisky, E., Paninski, L. (2016). **Multilayer Recurrent Network Models of Primate Retinal Ganglion Cell Responses**

Borst, A., Theunissen, F. (1999). **Information theory and neural coding.** 2(11), 947 - 957. https://dx.doi.org/10.1038/14731

Butts, D., Goldman, M. (2006)**. Tuning Curves, Neuronal Variability, and Sensory Coding** PLoS Biology 4(4), e92. https://dx.doi.org/10.1371/journal.pbio.0040092

Euler, T., Hausselt, S., Margolis, D., Breuninger, T., Castell, X., Detwiler, P., Denk, W. (2008). **Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina** Pflügers Archiv - European Journal of Physiology 457(6), 1393 - 1414. https://dx.doi.org/10.1007/s00424-008-0603-5

Franke, K., Berens, P., Schubert, T., Bethge, M., Euler, T., Baden, T. (2017). **Inhibition decorrelates visual feature representations in the inner retina** Nature Publishing Group 542(7642), 1 - 19. https://dx.doi.org/10.1038/nature21394

Litke, A., Bezayiff, N., Chichilnisky, E., Cunningham, W., Dabrowski, W., Grillo, A., Grivich, M., Grybos, P., Hottowy, P., Kachiguine, S., Kalmar, R., Mathieson, K., Petrusca, D., Rahman, M., Sher, A. (2004). **What does the eye tell the brain?: Development of a system for the large-scale recording of retinal output activity** Nuclear Science, IEEE Transactions on 51(4), 1434 - 1440. https://dx.doi.org/10.1109/tns.2004.832706

McIntosh, L., Maheswaranathan, N., Nayebi, A., Ganguli, S., Baccus, S. (2016). **Deep Learning Models of the Retinal Response to Natural Scenes.** Advances in neural information processing systems 29(), 1369 - 1377.