**Specific aims**

The retina is composed of multiple pathways, with each pathway processing specific visual information. These pathways do not process all visual information equally: some information is better represented by neurons than other, supposedly because it is more adaptive. While we know the physiologically behind such phenomenon, our understanding of why evolution designed the retina to better process certain types of information, is currently lacking. To truly understand what the retina does, we need mathematical theories that can explain as much of the retinal physiology with as few assumptions as possible. The efficient coding hypothesis is one of the most successful theories in this vein, which states that sensory systems should remove redundancies in their inputs to optimize the information they process. This hypothesis has been especially successful in the retina, where it can explain many features of retinal encoding such as center-surround receptive fields and ON-OFF pathways1, 2, 6, 9, 10. However, most of these models make simplifying assumptions that are biologically unrealistic, such as an infinite number of neurons and the possibility of negative firing rates. Because of these simplifying assumptions, it is currently unclear what predictions the efficient coding hypothesis makes about retinal pathways. To fill this gap in knowledge, I will use machine learning model to build a more flexible model that can make efficient coding predictions about the structure of neurons into different pathways. More specifically, I will provide efficient coding predictions about how the retina should process chromatic inputs (Aim 1), and about how the retina should process motion (Aim 2). By comparing these results with experimental data, we will know whether the efficient coding hypothesis is sufficient to explain how the retina encodes natural scenes.

**Aim 1: Expand efficient coding models to encompass chromatic information**

Encoding colors starts at the level of cone photoreceptors in the fovea of the retina. which come in three types –Long (L), Medium (M), and Short (S), roughly encoding red, green and blue stimuli, respectively11. This color information is ultimately encoded by RGCs, which are separated into different pathways. In macaques, midget cells represent roughly 80% of RGCs and encode the discrepancy between L and M cones, which tend to be highly correlated in natural images. Parasol cells (~10%) are separated into two pathways (ON and OFF) that sum L, M and S inputs. Bistratified cells (~10%) integrate ON inputs from S cones and OFF inputs from L and M cones. Bistratified cells are unique in that they asymmetric, and are the only major cell type to have an ON but not an OFF pathway. My model will try to explain two major findings: 1) Why this asymmetry in bistratified cells occurs, and 2) why midget cells represent most neurons despite encoding events that rarely occur in natural images. Completion of this aim will inform us whether efficient coding can successfully explain how the retina processes color in different pathways.

**Aim 2: Expand efficient coding models to explain motion-selectivity in RGCs**

Visual scenes are typically in motion, either because of objects moving or optic flow from our own movements. The encoding of visual motion starts as early as the retina, with several subtypes of retinal ganglion cells (RGCs) having stronger responses to one direction of motion than to its opposite16, 17. The major type of motion encoding RGCs are ON-OFF direction-selective ganglion cells (DSGC)18, which are divided into four subtypes that respond preferentially to each of the four cardinal directions17. There are currently no efficient coding predictions about these ON-OFF DSGCs. To fill this gap in knowledge, I will build an efficient coding model that jointly estimates receptive fields in space and time – a requirement to encode directional motion. I will then test whether we can replicate ON-OFF DSGCs that only encode the four cardinal directions17, and also compare my findings to experimental data from the Field Lab at UCLA. Completion of this aim will result in testable predictions for efficient coding of motion in natural images, which will help validate or invalidate the efficient coding hypothesis.