From Neurons to Gaze:
Continuous Decoding of Eye Movement Kinematics

1 Overview

Eye movements play an essential role in vision by directing our gaze toward points of interest. In terms of machine learning and brain function, eye movements offer a compelling model system for decoding various forms of brain activity, including attention, memory, decision-making, and more. However, deciphering the relationship between brain activity and the precise control of eye movements in real-time remains elusive, a challenge exacerbated by the inherent complexity of brain signals. Yet decoding eye movement-related brain activity is crucial not only for understanding the biological mechanisms of visual processing, but also for driving innovations in health and wellness applications. With modern eye tracking technology and increasingly powerful brain recording techniques, we can capture eye kinematics and neural activity, producing a rich source of time series data with high spatial and temporal resolution. This project provides an efficient pathway for harnessing ML techniques and deep learning architectures to uncover meaningful time-dependent information embedded within brain activity.

The central question to my research is: How can we effectively reconstruct eye movements from brain signals to interface with immersive devices and reveal meaningful patterns in physiological time series data? My investigation factors the problem into three pieces:

- **Leveraging AI models to decode eye trajectories from neural signals:** Using neural population data recorded from an awake, behaving monkey, we used supervised machine learning models to create a mapping between high-dimensional neural activity and eye movement behavior. I will further this work by overseeing experimental design and data recording, collecting a substantially richer dataset from the monkey as it performs a variety of complex tasks. This expanded dataset will focus on situations that require stabilizing gaze on unpredictably moving targets, and it will allow us to further refine the eye-brain mapping in service of a real-time cursor control on a screen.

- **Optimizing spatiotemporal resolution and model complexity for precise and robust real-time decoding:** My ongoing work involves assessing a range of supervised learning approaches, from elementary linear models to more sophisticated deep learning architectures, while exploring trade-offs between spatial and temporal resolution and their impact on predictive capabilities. In future work, I will assess the suitability of continuous versus discrete signals and employ dimensionality reduction techniques to enhance model interpretability within the context of natural primate vision, seeking insights into the factors that lead to robust decoding.

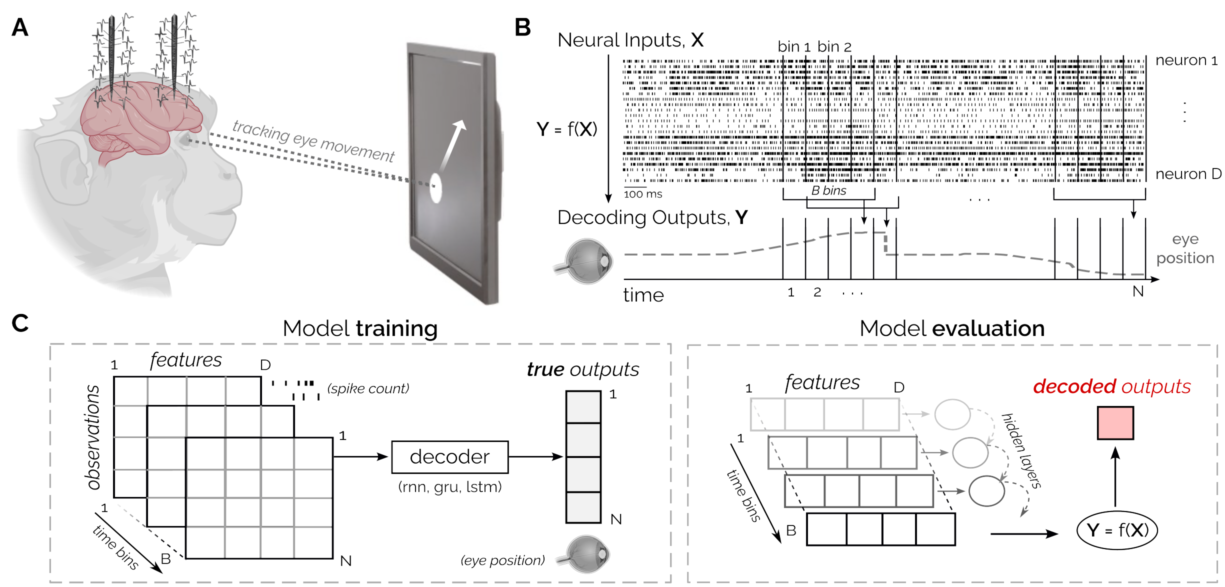
- **Domain adaptation for cross-species generalization:** Using eye tracking data I collected from human subjects, I will leverage the mapping established between monkey neural activity and eye tracking data to enhance the predictive accuracy of human eye tracking technologies, offering a non-invasive alternative to the need for invasive brain recordings.

2 Recent Research Results

During the last three years, I have researched how visual information that enters your eyes is propagated into your brain and, in turn, influences your eye muscles. Neuroscientists traditionally study the neural control of movements by analyzing firing rates averaged across many neurons. However, this traditional, static approach is limited because it cannot capture the dynamic interplay between neural activity and eye movements that occurs during natural vision. By harnessing the power of machine learning models to decode the intricacies of neural control over eye movements, we can not only illuminate the inner workings of visual-motor processing but also overcome a critical bottleneck in designing human-computer technologies that mimic the nuances of natural vision. Without a deeper understanding of how the brain encodes meaningful information about eye movements, the integration of these fundamental visual processes into such technologies remains a significant challenge that must be addressed for their effective implementation. Moreover, addressing this knowledge gap will facilitate the interpretation of eye tracking data, enhancing its utility in biomarker identification, neurodegenerative disease diagnoses, and cognitive tracking. **My work establishes a transformative approach to studying the connection between the brain and the eyes by employing machine learning to continuously decode eye movement trajectories from neural activity, offering a real-time measure of how the brain governs where we look.**

2.1 Leveraging AI models to decode eye trajectories from neural signals

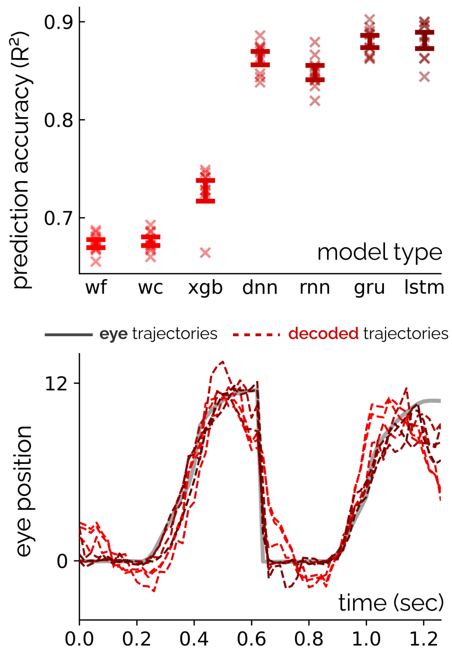
Continuous Mapping of Brain-Eye Dynamics with Tracking Eye Movements In this work, we employed classic decoding algorithms and neural network models to reconstruct eye movement kinematics from neural population responses recorded in the cerebral cortex of a monkey.



Neuron-to-behavior decoding pipeline. (A) Neural activity is recorded in cerebral cortex with multi-sensor probes while a monkey visually tracks a moving dot on a screen. (B) Neural spike timestamps are discretized into time bins and aligned with concurrently measured eye position. (C) Recurrent decoders are trained to reconstruct eye trajectories from sequential neural data, where each neuron's spike count serves as a distinct feature (N timepoints $\times$ D neurons, with temporal connections across B bins).

Tracking eye movements, characterized by the ability to follow a moving object with our gaze, offer an ideal window into understanding how the brain generates precise eye movements.

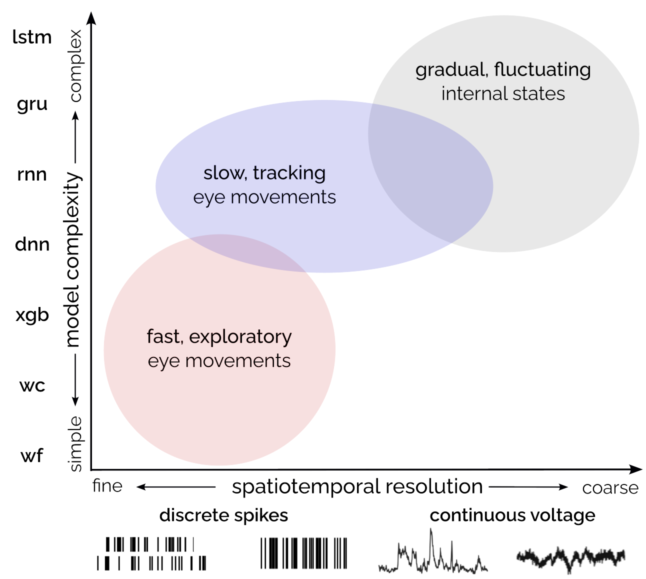
These movements involve a continuous adjustment of gaze direction to match a moving target, providing a testbed for interrogating the brain's ability to capture temporal dependencies and orchestrate movements on a moment-by-moment basis. Our results reveal that decoding eye position from relatively small populations of cortical neurons yielded impressive outcomes, explaining nearly 90% of the variance. Using a range of models, including simple linear filters, XGBoost, DNNs, vanilla RNNs, GRUs, and LSTM, we evaluated their performance in reproducing eye kinematics by quantitatively comparing the models using R^2 scores on held-out test sets. The high predictive power achieved through modern ML methods enables us to capture non-linear relationships within the data, decoding richer information content within the brain activity. Now, having established this first-of-its-kind neuron-to-behavior mapping for tracking eye movements, we can leverage example-based and model-agnostic techniques, including methods like Local Interpretable Model-Agnostic Explanations (LIME) and permutation feature importance, to further elucidate the information embedded within these neural signals and determine how cognitive factors, such as attention and arousal, exert influence on eye behavior.



Example decoding results. Prediction performance of seven model types and snippet of true and decoded y-positions.

2.2 Optimizing spatiotemporal resolution and model complexity for precise and robust real-time decoding

As we strive to achieve real-time gaze decoding for immersive visual devices, we must identify the optimal spatiotemporal resolution and granularity of the data that can enhance predictive performance, accommodate diverse behavioral outputs, and ensure computational efficiency. To address these issues, we explore different levels of temporal granularity and varying model complexities, investigating how they influence the quality and relevance of information extracted from the time series data.



Model framework for various output behaviors. Selection of decoding architecture (y-axis) and extracted neural features (x-axis), such as discrete single-unit or multi-unit activity or continuous voltage signals including entire-spiking activity or local field potentials.

Resolution, Granularity and Signal Quality of Neural Inputs This investigation uses various strategies, including assessing the optimal level of downsampling and leveraging continuous voltage signals (e.g., local field potentials) to reveal the content of eye movement-related brain activity across a range of timescales. While our neural probes and infrared eye tracker record at the 1-ms resolution, this level of resolution can introduce chronic instability, rendering the input signal unsuitable for real-time interfacing. In our preliminary work, we found that discrete, 50-ms binned signals yielded the highest accuracy in predicting moment-by-moment eye kinematics. Our findings also suggested that providing the decoding model with more neuronal data is crucial for improving model performance, but bifurcation of the input data across model features may not be the primary factor influencing the relevance of information for decoding eye movement trajectories. In our ongoing work, we are investigating which extracted neural features are best suited for decoding various cognitive aspects essential for natural vision, including exploratory "jumps" of the eyes and prolonged changes in mental states.

Model selection Another important design consideration that significantly impacts the robustness of gaze decoding is the choice of decoding algorithm. In our work, we benchmarked seven different decoding algorithms, each with their own assumptions on how neural inputs relate to behavioral outputs. While a few decoders consistently outperformed the others, surprisingly, the simple linear decoders demonstrated significant performance gains as well. However, these simpler models faced challenges when decoding specific nonlinear aspects of the behavior, such as tracking visual targets with varying speeds or capturing the monkey's responses to the onset of motion, which exhibited variable timings. To unlock the full potential of gaze decoding, models must account for a range of complex stimuli, accommodate various types of eye movements, and be able to autonomously determine which features within the neuronal data contain pertinent information at any given moment.

STILL BEING WRITTEN, THESE ARE JUST NOTES BELOW

3 Future Work

Over the next two years, I plan to build on my current results Recording more monkey and human data, making cross-species generalizations Incorporate saliency maps, pupil diameter Eye tracking offers a possibility to assess human and nonhuman primates using the same or similar conditions, which provides a translational opportunity to bridge the gap between basic and clinical investigations.

Predicting behavioral and cognitive states via real-time neural decoding: Test our mapping in realtime

Motor disabilities or conditions like ALShave limited mobility. High-res eye tracking can enable them to control computers and devices using their eye movements. Real-time feedback can enhance the accuracy and efficiency of assistive technology, improving the quality of life for individuals with disabilities. High-resolution eye tracking can enhance HCI by enabling precise control of devices through eye movements

Neurological health monitoring

Abnormal eye movement patterns can be indicative of neurological disorders such as Parkinson's disease or multiple sclerosis. High-resolution eye tracking can detect subtle changes that may occur before other symptoms.

We can record from many brain regions simultaneously in the monkey, and pinpoint where the locus of damage is coming from based on contributions of neural features to different aspects of behavior (pursuit v. saccades)

Real-time monitoring and feedback can help patients and healthcare providers track disease progression and adjust treatment plans as needed

The above mentioned future projects are necessary and complimentary for building

References