Title:

Cross Correlation of *in vivo* Neuropixel Recordings of Neuronal Responses During Evoked OKR in the Mouse Superior Colliculus

Author: John Gaynes

Abstract:

This project aims to explore the neuronal responses in the mouse superior colliculus (SC) during the optio-kinetic reflex (OKR) using *in vivo* Neuropixel recordings. By performing clustering analysis on local field potential (LFP) recordings from thousands of neurons, the study seeks to identify and characterize functional neuron ensembles involved in the OKR and their integration with visual inputs. The project includes a comprehensive approach to data collection, preprocessing, and clustering analysis, with the goal of identifying patterns of neuronal activity that underlie this sensory-motor reflex. The findings will facilitate understanding how the SC processes sensory information and coordinates motor outputs, particularly in the context of visual tracking and movement detection.

Introduction

The superior colliculus (SC) is an important structure in the midbrain that plays an essential role in integrating sensory information, particularly visual inputs, and coordinating motor responses, such as saccadic eye movements. One of the key functions of the SC is to moderate the optio-kinetic reflex (OKR), an involuntary eye movement that allows humans and animals to track and stabilize images on the retina as they move through the surrounding environment. This reflex is enables animals to maintain visual stability and accurately track moving objects.

This project seeks to answer several important questions about the neural mechanisms underlying OKR, focusing on the functional connectivity and organization of neurons within the SC. Specifically, the research questions addressed are:

1. **How are neurons in the superior colliculus functionally connected during the optio-kinetic reflex?**

Understanding the functional connectivity between neurons is essential to understand how the SC processes visual information and coordinates motor responses. By examining the relationships between neurons, we can identify potential neuron ensembles that work together to execute the reflex.

2. **Can distinct neuron ensembles be identified based on their neuronal waveform characteristics?**

Neuronal waveforms, which include attributes such as amplitude, duration, rise time, and decay time, are critical indicators of a neuron's activity and function. Identifying distinct ensembles based on these waveform characteristics could reveal how different groups of neurons contribute to the overall functionality of the SC.

3. **What patterns of neuronal activity contribute to the sensory-motor integration within the SC during OKR?**

By analyzing the patterns of activity across thousands of neurons, this project aims to uncover the specific neural dynamics that enable the SC to integrate sensory inputs and produce coordinated motor outputs. These patterns are likely key to understanding how the SC achieves its role in sensory-motor integration.

These are important questions to understand since the SC plays an essential role in the brain's network for visual processing and motor coordination. The insights gained from this study could provide an increased understanding of the SC's role in visual-motor integration, which has implications for a wide variety of neurological conditions that affect vision and movement, such as strabismus, amblyopia, and other disorders. Furthermore, understanding the organization and connectivity of neurons within the SC could inform the development of new therapeutic approaches and technologies, such as brain-computer interfaces, aimed at restoring or enhancing visual-motor functions.

Related Work:

The dataset used in this project is novel and unpublished, providing a unique opportunity to study the SC's role in OKR with a high level of detail. Although there is no directly related previous work on the specific dataset, existing literature provides foundational insights into the general methodologies for recording and analyzing neural activity in the SC.

1. **Sibille, J., Gehr, C., The, K.L., Kremkow, J. (2022).** "Tangential high-density electrode insertions allow to simultaneously measure neuronal activity across an extended region of the visual field in mouse superior colliculus." *Journal of Neuroscience Methods, Volume 378.*

This paper outlines the method for using high-density electrodes to record neuronal activity in the SC. Although it does not specifically analyze neural activity during evoked OKR, it provides a basis for understanding the technical approaches used to collect this data.

In addition to this study, other studies have explored various aspects of SC function, particularly its involvement in saccadic eye movements and visual processing. Previous research has established the SC as a crucial hub for integrating sensory inputs and coordinating motor responses, especially in response to visual stimuli. The current project builds on these studies by applying advanced data analysis techniques to Neuropixel recordings, aiming to gain insight into the functional organization of the SC during OKR.

Dataset

The dataset used in this project is a rich and complex collection of neuronal recordings obtained from the superior colliculus (SC) of live mice. This data was gathered during experiments designed to evoke the optio-kinetic reflex (OKR), a critical reflexive eye movement that allows the mouse to maintain visual stability. The dataset is unique, offering a detailed view of the SC's activity at a large scale and high resolution.

Source of the Data

The dataset originates from a series of 133 experiments conducted on live mice, each involving the implantation of a Neuropixel probe into the SC. Neuropixel probes are advanced electrophysiological tools capable of recording from hundreds of neurons simultaneously across multiple brain regions. In this project, each probe featured 386 channels, enabling the recording of neuronal activity from a vast array of neurons in the SC.

Given the depth and scope of these recordings, the dataset provides an extensive resource for analyzing the SC's role in sensory-motor integration during OKR. The data is stored securely on a private Google Drive and has been backed up on two PCs.

Attribute Features

The dataset is highly detailed, capturing several key attributes of the neuronal waveforms recorded during the experiments. These attributes include:

Amplitude: This represents the peak value of the neuronal waveform, which is indicative of the strength of the neuronal response to stimuli. It is an important measure of neural activity during the reflex.

Duration: This is the time span of the neuronal waveform, measuring the length of time the neuron remains active during a response. Duration helps in understanding the temporal kinetics of neuronal activity.

Rise Time: This is the time taken for the waveform to reach its peak amplitude from the baseline. It reflects the speed at which a neuron responds to stimuli, providing insights into the neuron's excitability and reaction time.

Decay Time: This measures the time it takes for the waveform to return to baseline after reaching the peak. Decay time is important for understanding how long a neuron remains active and how quickly it recovers after firing.

These attributes were chosen for their relevance to the study's objectives, as they provide essential information about the characteristics and behaviors of individual neurons. By analyzing these attributes, researchers can gain insights into the functional connectivity and organization of neurons within the SC.

Data Collection Process

The data collection process involved several meticulous steps to ensure the accuracy and reliability of the recordings:

1. **Neuropixel Probe Implantation:** Neuropixel probes were carefully implanted into the SC of each mouse, positioned to capture neuronal activity across a broad area of the SC. The probes' high-density design allowed for the simultaneous recording of activity from multiple neurons, providing a comprehensive view of the SC's response during OKR.

2. **Spike Sorting:** Once the data was collected, it was preprocessed using Kilosort, a sophisticated spike-sorting algorithm. Kilosort identifies and classifies neuronal spikes within the data, differentiating between individual neurons and organizing the data for further analysis.

3. **Data Organization:** The spike-sorted data was then compiled into a single, comprehensive CSV file. This file was organized by animal, recording session, and individual neuron, ensuring that the data was organized for detailed analysis.

Given the large volume of data generated by these experiments, preprocessing and organization were critical to enable effective analysis. The detailed nature of the dataset, combined with the advanced techniques used in its collection and preparation, make it a valuable resource for studying the SC's functional architecture and its role in OKR.

Confidentiality and Access

Due to the sensitive nature of the data and its importance in ongoing research, the dataset is not publicly accessible. It is stored securely and is only available to authorized researchers involved in this project. This ensures that the data remains confidential while still allowing for in-depth analysis and exploration by the project team.

The dataset provides a unique and powerful foundation for this project, enabling the interrogation of neuronal activity in the SC with a level of detail that was previously unattainable. The insights gained from this data have the potential to significantly advance our understanding of the SC's role in sensory-motor integration and reflexive eye movements.

Main Techniques Applied

This section outlines the main techniques applied throughout the project, encompassing data cleaning and preprocessing, data organization and storage, and the advanced classification and clustering methods used to analyze the neuronal recordings from the superior colliculus (SC) during the optio-kinetic reflex (OKR).

Data Cleaning and Preprocessing

Given the complexity and volume of the data collected from the Neuropixel probes, meticulous data cleaning and preprocessing were essential to ensure the accuracy and reliability of the analysis. The data cleaning process involved several key steps:

1. **Handling Missing Values:**

The initial step in data preprocessing was to identify and handle missing values, particularly NaNs, which can disrupt analyses if left untreated. An automated script was used to scan the dataset for NaNs, and visual inspections were conducted to verify the results. NaNs were replaced with zeros, a decision made to maintain the continuity of the data without introducing biases that might arise from more aggressive imputation methods.

2. **Noise Reduction:**

Neuronal recordings often contain noise, which can obscure the biological signal. Filtering techniques were used to reduce noise levels and enhance the clarity of the waveforms. This step was critical for ensuring that the subsequent analysis focused on genuine neuronal activity rather than artifacts.

3. **Spike Sorting:**

Kilosort, a widely used spike-sorting software, was used to process the raw data from the Neuropixel probes. Kilosort identifies and classifies neuronal spikes, distinguishing between individual neurons and organizing the data for further analysis. This process was crucial for converting the raw recordings into a format suitable for clustering and correlation analysis.

4. **Data Normalization:**

To facilitate comparison across neurons and recording sessions, the data was normalized. This involved adjusting the values within the dataset to ensure that all waveform attributes (such as amplitude and duration) were on a comparable scale. Normalization was necessary to prevent any single attribute from disproportionately influencing the clustering results.

Data Organization and Storage

The sheer volume of data generated by the 133 experiments required a robust data organization and storage strategy. The following steps were implemented to manage the dataset effectively:

1. **Data Compilation and Structuring:**

After preprocessing, the spike-sorted data was compiled into a CSV file. This file was organized by animal, recording session, and neuron, creating a structured dataset that facilitated efficient access and analysis. Each row represented a neuron, and each column corresponded to a specific attribute or condition.

Data Storage:

The dataset was stored in a secure data warehouse on a private Google Drive, with redundancy measures in place through backups on two PCs. This setup ensured the data's safety and integrity, allowing for secure and reliable access throughout the project. The use of hierarchical data formats (HDF5) within the dataset further enhanced the ability to handle large volumes of data efficiently, enabling quick retrieval and manipulation of specific subsets of the dataset.

2. **Data Access and Management:**

To manage and explore the hierarchical data formats, HDFView software was employed. This tool facilitated the inspection and manipulation of the large datasets, allowing for efficient data management and enabling the project team to focus on the analysis rather than data handling.

Classification and Clustering Techniques

The core of the analytical approach in this project involved advanced classification and clustering techniques, which were applied to identify and characterize neuron ensembles within the SC. The following methods were central to this analysis:

1. **K-means Clustering:**

K-means clustering was used to group neurons based on their waveform attributes. This unsupervised learning technique partitions the data into a predefined number of clusters, in this case, eight. Each neuron was assigned to the cluster that minimized the variance within the cluster, resulting in distinct groups of neurons with similar characteristics. K-means was chosen for its simplicity and effectiveness in handling large datasets, and it provided a clear separation of neuron ensembles.

2. **Hierarchical Clustering:**

In addition to K-means, hierarchical clustering was applied to explore the relationships between neurons at multiple levels. Using Ward's method, the MATLAB function constructed a tree-like structure that revealed the hierarchical relationships among the neurons. This method allowed for the identification of sub-clusters within the main clusters, offering a more nuanced understanding of the functional organization of neurons within the SC.

3. **Principal Component Analysis (PCA)**:

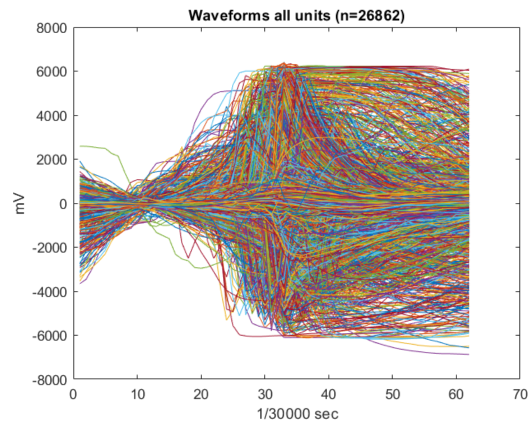
To reduce the dimensionality of the data and make the clustering results more interpretable, Principal Component Analysis (PCA) was employed. PCA transforms the original data into a new set of uncorrelated variables (principal components), which capture the most significant variance in the dataset. By focusing on the first few principal components, the project was able to visualize the clusters and gain insights into the key features that distinguish different neuron ensembles.

4. **Scatter Plot Analysis and Regression**:

To further explore the relationships between waveform attributes (such as Min amplitude, Max amplitude, and AUC), scatter plots were generated, and regression analysis was performed. This step provided additional insights into how these attributes relate to each other within and across clusters, helping to identify potential functional relationships between neurons.

These classification and clustering techniques were crucial for uncovering the functional organization of the SC and understanding how different neurons contribute to the optio-kinetic reflex. The combination of K-means and hierarchical clustering, supported by PCA and scatter plot analysis, provided a multi-faceted approach for analyzing the neuronal recordings and identifying meaningful patterns within the data.

Key Results

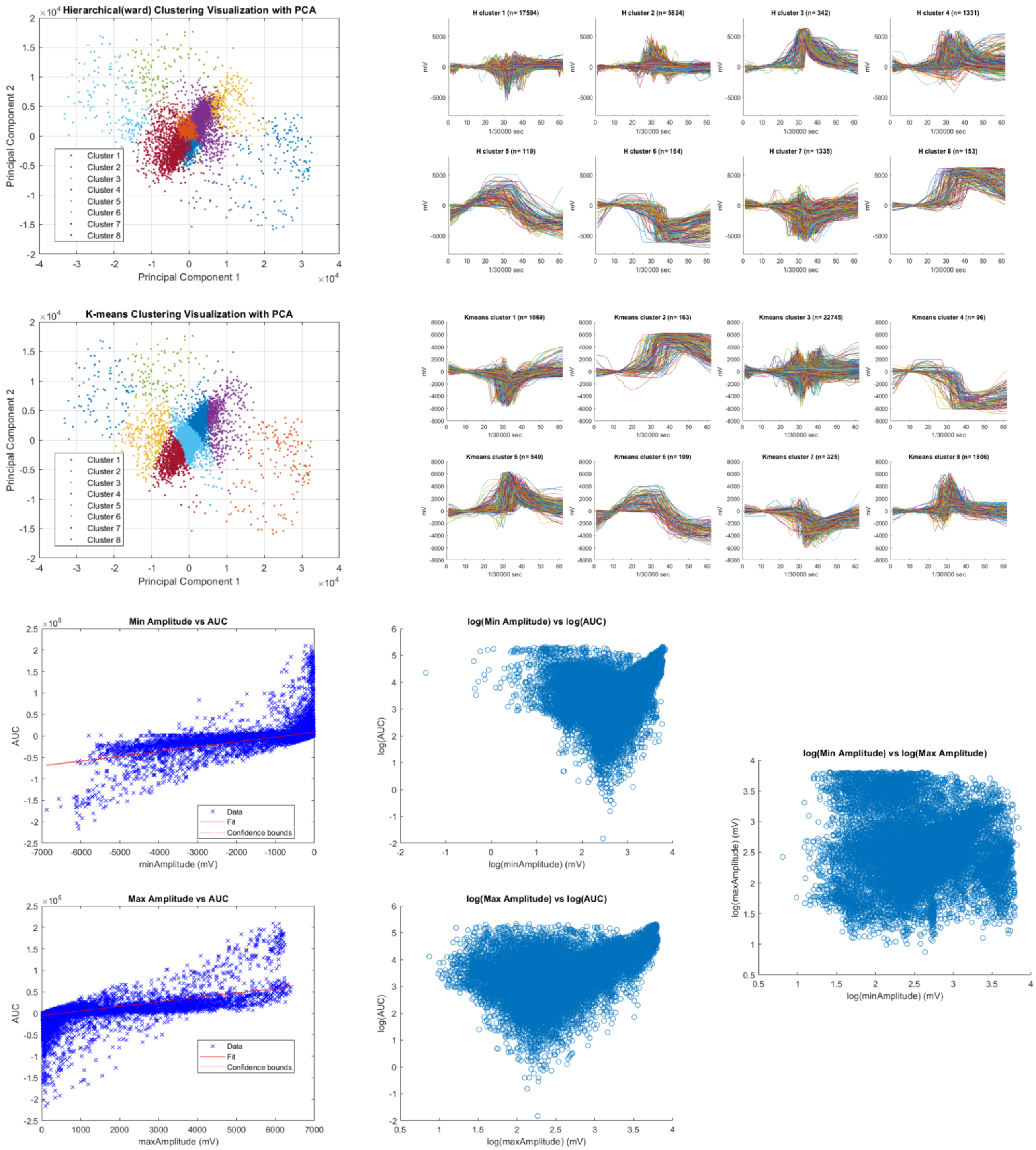


The analysis of the neuronal recordings from the superior colliculus (SC) during the optio-kinetic reflex (OKR) yielded several key findings, providing valuable insights into the functional organization and connectivity of neurons within this critical brain region. Through the application of advanced clustering techniques and detailed waveform analysis, the project uncovered distinct patterns of neuronal activity that contribute to the SC's role in sensory-motor integration.

1. Identification of Distinct Neuron Ensembles

One of the most significant findings was the successful identification of distinct neuron ensembles within the SC, achieved through

both K-means and hierarchical clustering. The clustering analysis grouped neurons based on their waveform attributes—such as amplitude, duration, rise time, and decay time—revealing eight well-defined clusters. Each cluster represented a group of neurons with similar waveform characteristics, suggesting that these neurons may share



common functional roles in processing visual inputs and mediating the OKR.

The clear separation between clusters indicated that the SC is organized into functionally distinct compartments, with each potentially specialized in processing different aspects of visual information or coordinating specific types of motor responses. This discovery enhances our understanding of how the SC integrates sensory inputs and produces coordinated motor outputs, a process that is essential for maintaining visual stability during movement.

2. Functional Connectivity Insights

The project also provided new insights into the functional connectivity of neurons within the SC. The cross-correlation analysis revealed that neurons within the same cluster were more likely to be functionally connected, as indicated by the strong correlations between their activity patterns. This finding supports the hypothesis that the SC is not just a collection of independent neurons but rather a network of interconnected neurons that work together to achieve sensory-motor integration.

In addition, the hierarchical clustering analysis suggested that these functional connections are organized in a hierarchical manner, with some neurons playing more central roles in the network than others. This hierarchical organization could reflect the SC's ability to prioritize certain visual inputs or motor commands, ensuring that the most critical information is processed and acted upon quickly.

3. Relationships Between Waveform Attributes

Further exploration of the relationships between waveform attributes—such as Min amplitude, Max amplitude, and the area under the curve (AUC)—revealed additional layers of complexity within the SC's functional organization. Scatter plots and regression analysis showed that some neurons exhibited positive correlations between these attributes, indicating that neurons with more pronounced waveforms tend to have larger overall responses.

The log-transformed scatter plots provided similar insights, confirming that while many neurons showed coordinated responses across different attributes, others did not. This variation suggests a level of heterogeneity in how neurons within the SC process visual information and contribute to the OKR. The diversity in these relationships demonstrates the complexity of the SC's role in sensory-motor integration and suggests that different neuron ensembles may be specialized for processing different types of visual stimuli.

4. Robustness and Interpretability of Clustering Results

The use of Principal Component Analysis (PCA) to support the clustering was effective in reducing the dimensionality of the data and enhancing the interpretability of the results. By focusing on the most significant features captured by the principal components, the project was able to visualize the clusters in a way that clearly distinguished between different neuron ensembles. This approach not only validated the clustering results but also provided a deeper understanding of the key features that define each cluster.

Overall, the combination of K-means and hierarchical clustering, supported by PCA and detailed waveform analysis, allowed for a comprehensive exploration of the SC's functional architecture. The key results from this project contribute to a deeper understanding of how the SC processes sensory information and coordinates motor outputs, particularly in the context of the OKR. These findings have broader implications for understanding the neural mechanisms underlying sensory-motor integration and could inform future research on related topics in neuroscience.

Applications

The findings from this project have significant implications and potential applications across several fields, ranging from basic neuroscience research to clinical interventions and the development of advanced neurotechnological tools. By uncovering the functional organization and connectivity of neurons within the superior colliculus (SC), this research contributes to a deeper understanding of sensory-motor integration, particularly in the context of the optio-kinetic reflex (OKR). The insights gained can be applied in various ways:

1. **Advancing Neuroscience Research**

The detailed characterization of neuron ensembles within the SC provides a valuable framework for further research on sensory-motor integration. By identifying distinct groups of neurons that are functionally connected, this study lays the groundwork for exploring how these ensembles interact with other brain regions involved in visual processing and motor control. Future research can build on these findings to investigate how the SC integrates inputs from the retina and other sensory modalities, coordinates complex motor responses, and adapts to changes in the environment.

Additionally, the methodologies developed in this project, including the use of Neuropixel probes, advanced clustering techniques, and waveform analysis, can be applied to study other brain regions and behaviors. Researchers can leverage these tools to investigate neural dynamics in different contexts, potentially leading to new discoveries about how the brain processes sensory information and generates coordinated outputs.

2. **Clinical Implications and Neurological Disorders**

The knowledge gained from this study has important clinical applications, particularly in understanding and treating neurological disorders that affect vision and motor coordination. Conditions such as strabismus, amblyopia, and Parkinson's disease involve disruptions in the normal functioning of the SC and related neural circuits. By characterizing the functional organization of the SC, this research provides a foundation for developing targeted interventions that can restore or enhance sensory-motor integration in patients with these conditions.

For example, therapies aimed at modulating specific neuron ensembles identified in this study could be developed to improve eye movement control in patients with strabismus. Similarly, understanding the neural basis of OKR could inform rehabilitation strategies for individuals with impaired visual tracking due to brain injuries or neurodegenerative diseases.

3. **Development of Brain-Computer Interfaces (BCIs)**

The insights into the SC's functional connectivity and neuron ensemble organization have potential applications in the development of brain-computer interfaces (BCIs). BCIs are technologies that enable direct communication between the brain and external devices, which are developed to be used to restore lost motor functions in individuals with severe disabilities. By mapping the neural circuits involved in sensory-motor integration, this research could facilitate the design of BCIs that enable or enhance the brain's natural processing of visual information and coordination of motor responses.

BCIs that integrate with the SC could be designed to assist individuals with visual impairments by enhancing their ability to track and respond to moving objects. Such interfaces could also be used in advanced prosthetics, where real-time visual feedback is crucial for controlling artificial limbs or other assistive devices.

4. **Enhancing Machine Learning and Artificial Intelligence Models**

The clustering techniques and analytical approaches used in this study can be applied to improve machine learning and artificial intelligence (AI) models, particularly those related to pattern recognition and classification. The successful identification of neuron ensembles based on waveform attributes demonstrates the effectiveness of unsupervised learning methods, such as K-means and hierarchical clustering, in idetifying unknown patterns in complex datasets.

These techniques can be adapted for use in various AI applications, such as image and speech recognition, where the ability to classify and interpret high-dimensional data is essential. By incorporating insights from neural data, AI models can be designed to more closely mimic the brain's natural processing abilities, leading to more accurate and efficient algorithms.

5. **Informing Neurotechnological Innovations**

Finally, the findings from this project have the potential to lead to new innovations in neurotechnology. The detailed understanding of how the SC processes sensory inputs and coordinates motor responses can be used to the develop new tools and devices for studying and manipulating brain activity. For example, neuroprosthetics that interface with the SC could be developed to restore vision or enhance motor control in individuals with neurological impairments.

The project's use of advanced recording and analysis techniques could be applied to the development of new neuroimaging tools that provide real-time insights into brain function. Such tools could be used in both research and clinical settings to diagnose and manage neurological disorders, determine the effectiveness of treatments, and explore the neural basis of behavior.