Neuronal Cell-Typing And Classification Using Human Single Neuron Recordings In Parkinson's Disease Patients

Jazlynn Xiu Min Tan, Regina Annirood, Georgia Bardaklis, & Xinyi Zhang
Project 5

University of Toronto
Toronto, Ontario, Canada

Abstract

Parkinson's disease (PD) profoundly affects quality of life, characterized by motor symptoms stemming from dopamine-producing neuron degeneration in the brain's substantia nigra. Deep Brain Stimulation (DBS) offers substantial relief when drug therapies falter. Precise electrode placement within the subthalamic nucleus (STN) and globus pallidus (GPi) is critical for DBS success. Imaging and microelectrode recordings guide surgical navigation, but manual recordings present challenges of subjectivity and variability.Here, we present a study involving 360 neurons from PD patients' STN and GPi. Using machine learning on spike train features, our model accurately classified STN/SNr and BOR/HFD neurons with 72-100% accuracy. We discuss the potential of simultaneous multi-neuron recordings to capture broader neural dynamics and lateralization patterns, offering enhanced insight into PD pathology.

Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder that has been presented in over 100,000 Canadians with almost 7000 new cases being diagnosed each year[1]. PD is characterized by the degeneration of dopamine-producing neurons in the substantia nigra region of the brain and this depletion of dopamine leads to a range of motor symptoms, including tremor, bradykinesia, rigidity, and postural instability[2,3]. As PD progresses over time, these symptoms worsen, often resulting in a decline in motor function and a substantial reduction in the quality of life for those affected.

Deep Brain Stimulation (DBS) has emerged as a key therapeutic option for managing the motor symptoms of PD when the efficacy of oral dopaminergic drugs begin to diminish.. DBS involves the accurate placement of surgically implanted electrode(s) into specific brain regions followed by the delivery of controlled electrical impulses[4]. The primary aim of DBS is to modulate abnormal neural activity and alleviate motor symptoms effectively. DBS has been shown to provide substantial and sustained improvement in motor function, making it a valuable treatment option for PD patients[5].

The precise implantation of electrodes is crucial for the success of DBS in treating PD. Common target sites in the dysfunctional circuits of the basal ganglia include the subthalamic nucleus (STN), the globus pallidus interna (GPi), and the ventral intermediate nucleus of the motor thalamus[5–7].

During electrode implantation, neurosurgeons encounter various anatomical structures along each trajectory. In the case of targeting the STN, the trajectory involves passing through structures such as the substantia nigra (SN) and the zona incerta. The SN has distinct segments, including the pars compacta (SNc) and pars reticulata (SNr)[8]. Identifying the borders between these segments is critical for precise electrode placement. The STN is a small, deep-seated nucleus adjacent to the SNr. Accurate localization of the STN-SNr border is essential to avoid unintended stimulation of the SNr, which could lead to undesirable side effects such as speech or cognitive disturbances[8,9]. Advances in imaging techniques, such as high-resolution MRI and diffusion tensor imaging (DTI), contribute to improved visualization and identification of these borders during surgery.

Similarly, targeting the GPi involves traversing through the internal capsule and the external globus pallidus (GPe)[10]. The internal capsule contains descending motor fibers, and precise electrode placement requires navigating through this structure. As with the STN, identifying the borders between the GPi and adjacent structures is crucial. The border between the GPi and the GPe is particularly important. High-frequency discharge (HFD) neurons within the GPi and GPi border cells (BOR) play significant roles in motor function regulation[10]. HFD neurons are characteristic of the GPi, and their identification during electrode implantation is vital for achieving therapeutic effects in PD. The GPi-HFD neurons exhibit increased firing rates during movement, making them a specific target for effective symptom modulation. The identification of GPi-HFD neurons relies on microelectrode recordings, emphasizing the importance of neurophysiological guidance during surgery[10–12]. GPi-BOR cells are located at the border between the GPi and GPe. These cells are important for distinguishing between the two structures. Accurate placement of electrodes within the GPi while avoiding the GPe is essential to prevent unintended stimulation of the GPe, which could lead to ineffective symptom control or unwanted side effects. Microelectrode

recordings play a key role in differentiating between GPi-BOR cells and neurons within the GPe, ensuring precise electrode placement within the GPi[11,12].

The DBS procedure involves several key steps, with MRI-based imaging, specialized neuronavigation planning software and microelectrode recordings playing pivotal roles. Imaging, typically MRIs aids in preoperative planning and target localization, however; relying solely on imaging has limitations, as it may not provide sufficient resolution to precisely identify small target structures. This is where microelectrode recordings become essential[4].

Microelectrode recordings involve the use of fine electrodes to record neural activity as they are gradually inserted into the brain. This technique allows clinicians real-time monitoring of neuronal firing patterns, helping to identify the optimal implantation sites. Neurophysiological guidance, combining imaging and microelectrode recordings at the level of a single neuron or local neuron populations through local field potential (LFP) analyses, ensures accurate targeting and improves the overall efficacy of DBS[4,13,14].

Despite advancements, manual microelectrode recordings during surgery pose challenges, such as subjectivity and intraoperative variability. This has led to the development of automated procedures to enhance precision and streamline the process. Automated methods, as explored by Contarino et al. (2014), involve algorithms that analyze microelectrode recordings to identify characteristic neuronal patterns, aiding in the determination of optimal implantation sites[15]. These automated procedures aim to reduce variability, improve accuracy, and enhance the overall efficiency of DBS surgeries.

In this study, we isolated the spike trains of 360 neurons recorded from the STN, SNr, HFD and BOR during awake neurosurgery of PD patients. We built a machine learning model to classify neurons in the subthalamic nucleus as STN or SNr, and neurons in the globus pallidus as BOR or HFD, using spike train features. Our methodology is done to predict the brain region where the neuron is located. For model validation, we will use cross-validation techniques to ensure robustness and generalizability. Our methodology aims not only to predict the neuronal origin but also to provide a deeper understanding of the functional interplay regions implicated in PD pathology.

Methods

Data Acquisition and Preprocessing

Extracellular recordings were extracted from 360 neurons in the subthalamic nucleus and globus pallidus of awake neurosurgical patients diagnosed with Parkinson's disease (Fig. 1D), consisting of STN, SNr, BOR and HFD neuron types. Single-neurons were sampled along the trajectory targeting either the STN or GPi as part of the standard neurophysiological mapping procedure[4,8,10]. Recordings were acquired at a sampling frequency of 12500 Hz. To isolate the frequency band for action potentials and remove high-frequency noise, the recordings were filtered with a bandpass filter and cutoff frequencies were set at 300 Hz and 6000 Hz.

Spike Sorting

Fig. 1A summarizes our spike sorting pipeline. Putative spikes were detected via a negative threshold crossing set as the 70th percentile of the filtered signal. Voltage traces of putative spikes were extracted from a window of 4 ms centered on the spike. Principal Component Analysis (PCA) was performed to reduce the waveform features to three principal components which were then clustered via K-Means clustering. The optimal number of clusters from a range of 2-5 clusters was identified by maximizing the silhouette score.

The mean waveforms of the potential spikes for each cluster were systematically visualized. Clusters were merged or discarded based on waveform shape, the number of detected spikes, signal-to-noise ratio (SNR), and refractory period (1ms) violations. The isolated neurons have a mean SNR of 2.84 (Fig. 1B) and an ISI isolation index of 1.37 (Fig. 1C). Spike trains were extracted from the curated clusters.

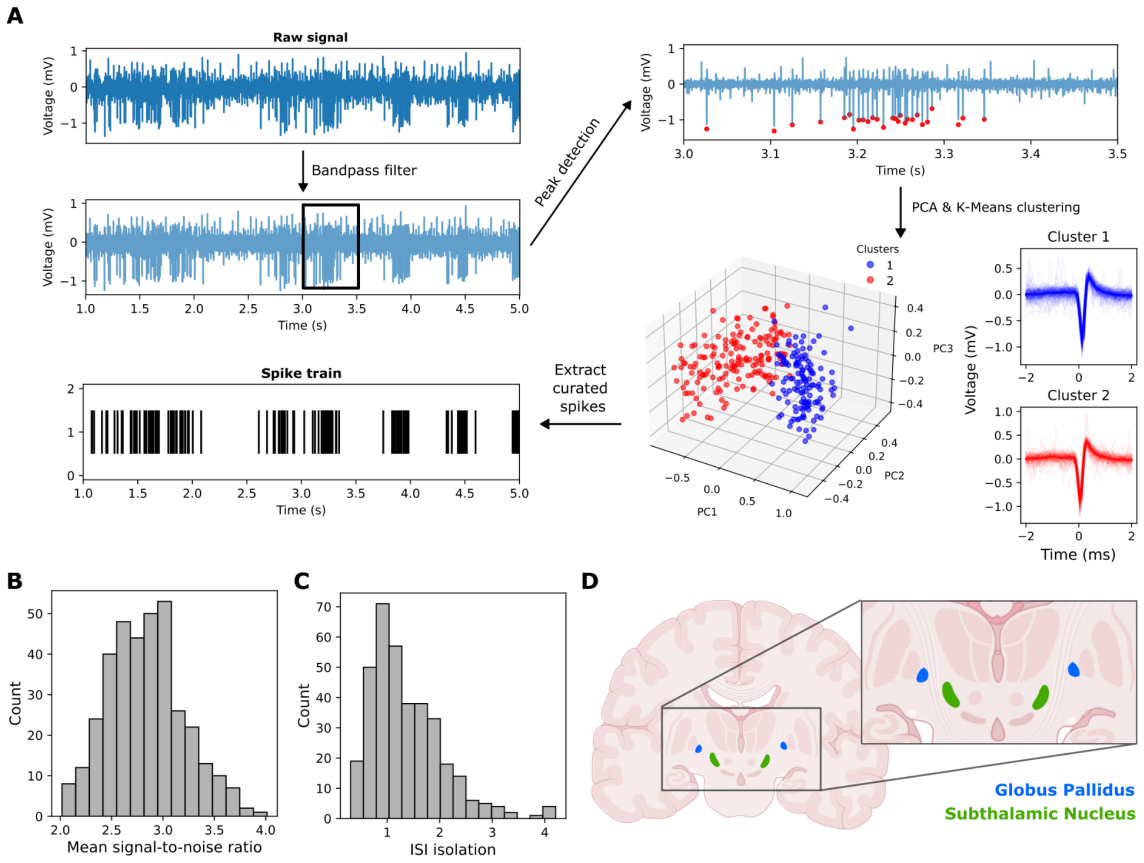


Figure 1: Spike sorting. (A) The raw signal is bandpass filtered prior to peak detection. 4ms of the filtered signal centered on the putative spike is extracted and undergo dimension reduction via PCA. The first 3 PC components are used in K-Means clustering where the optimal number of clusters is determined through silhouette analysis. Clusters are curated, merged or discarded to isolate single neurons. The spike train is then extracted from the curated spikes. **(B)** Isolated neurons have mean SNR of 2.84 and **(C)** ISI isolation index of 1.37. **(D)** The dataset comprises neurons from the subthalamic nucleus and globus pallidus.

Feature Engineering and Selection

Following spike sorting, we extracted 21 features (Table 1) that have had prior success in neuronal classification or have been described in PD pathology[16–18]. Instantaneous firing rate was calculated for a moving window of 100ms with a step size of 50ms. Burst activity was detected by the Poisson Surprise method and burst index was calculated as the ratio of mean inter-spike interval (ISI) to median ISI. Oscillatory activity was detected with the method described by Kaneoke & Vitek (1996)[19] where the autocorrelogram was calculated with a time lag of 500ms and a bin size of 10ms, and only frequencies meeting the significance threshold of $p < 0.05$ were retained.

Model training

We trained separate models to classify neurons in the STN as SNr or STN and neurons in the GPi and BOR or HFD, in the left and right hemispheres separately. The training set consisted of 80% of the data while 20% was held out as a test set. We first trained a random forest and ranked all features by Gini importance to select the top 4 for each model. Hyperparameter optimization was conducted with 5-fold cross validation. A logistic regression model served as a baseline and compared with random forest and SVM models. Prior to SVM training, all data was preprocessed with standard scaling. All models were evaluated on accuracy, defined as the percentage of correct predictions out of all predictions made.

Code availability

All code is available at our Github Repository.

Results

Random Forest Classification of STN and GPi Neurons

For the right subthalamic nucleus, the four features with highest Gini importance are: 1) mean interspike interval; 2) mean instantaneous firing rate; 3) mean surprise of bursts; 4) variance of burst duration (Fig. 2B, selected in pink). The kernel density estimation plots of these features show strong separation between the STN and SNr neurons (Fig. 2A). The logistic regression and SVM achieved 89% accuracy while the random forest achieved 94% accuracy on the test set (Table 2). The random forest misclassified one SNr neuron as an STN neuron as shown in the confusion matrix (Fig. 2C). For the left subthalamic nucleus, all models achieved 72% accuracy on the test set (Table 2, Supp Fig. 1A-C).

The three most important features are shared with the model for the right globus pallidus, with the addition of the number of bursts instead of the variance of burst duration as the fourth feature (Fig. 2E). There is strong separation between the BOR and HFD neurons for the selected features (Fig. 2D). The logistic regression and SVM achieved 94% and 89% accuracy on the test set respectively (Table 2). The random forest was able to correctly classify all neurons (Fig. 2C), achieving a 100% accuracy (Table 2). For the left globus

pallidus, the logistic regression, SVM and random forest achieved 78%, 83% and 89% accuracy respectively (Table 2, Supp Fig. 1D-F).

Pooling data from both hemispheres

To explore whether the GPi and STN of the two hemispheres had similar spike train features, we trained a random forest model using both hemispheres. The model was able to classify STN and SNr neurons in the STN with 95% accuracy on the test set (Table 3), which was an improvement from training on the hemispheres separately. In the GPi, using both hemispheres resulted in a lower test set accuracy of 89% (Table 3) compared to training on hemispheres separately.

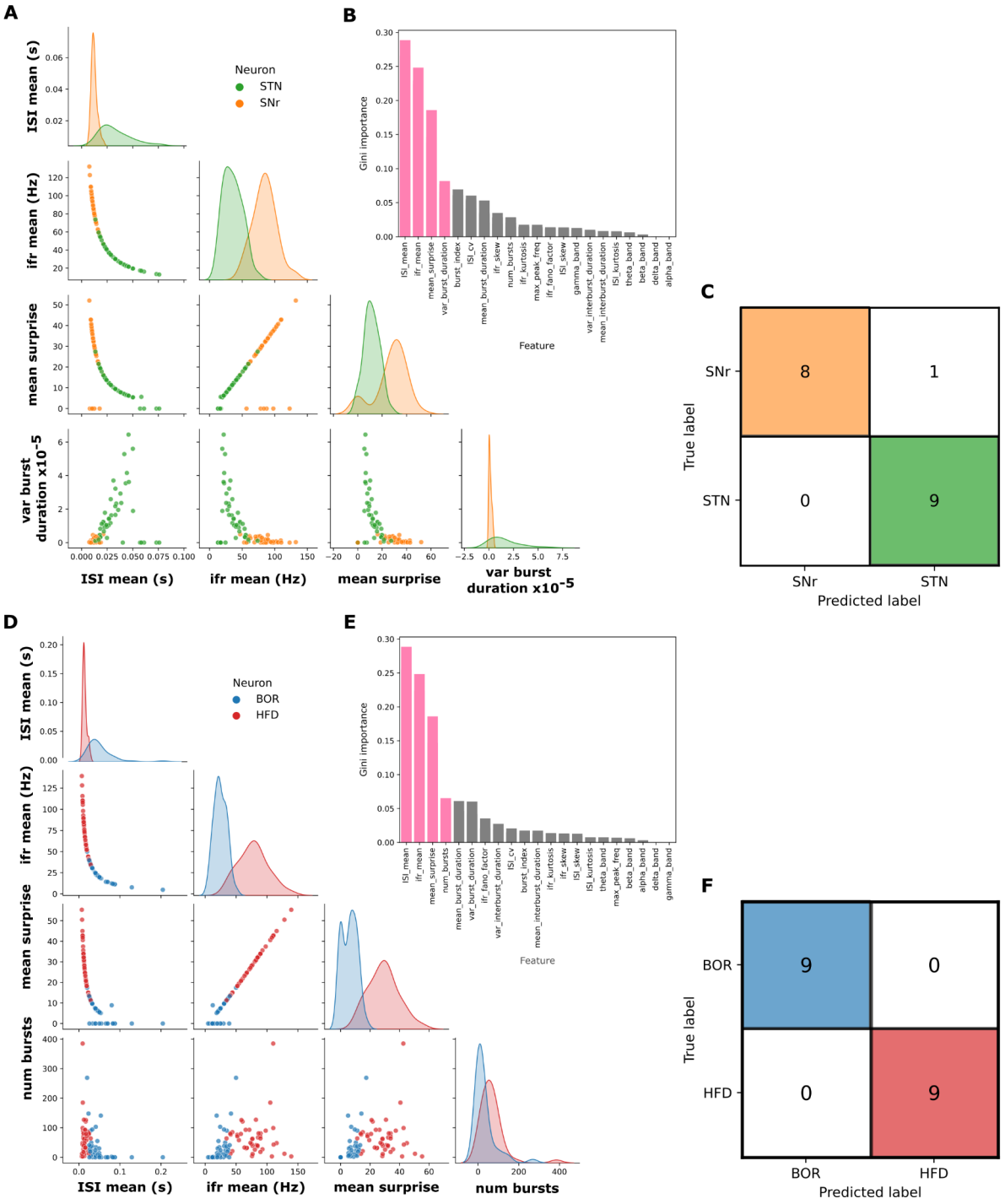


Figure 2: Classification of Neuron type in subthalamic nucleus and globus pallidus of the right hemisphere. (A) The four selected features with highest Gini importance (pink) **(B)** provide good separation between the STN and SNr neurons in the subthalamic nucleus. **(C)** On the test set, the random forest model achieves 94% accuracy, with only 1 SNr neuron misclassified as a STN neuron. **(D)** The four selected features with highest Gini importance (pink) **(E)** provide good separation between the HFD and BOR neurons in the globus pallidus. **(F)** On the test set, the random forest model achieves 100% accuracy.

Discussion

We trained random forest models capable of distinguishing STN and SNr neurons in the subthalamic nucleus and BOR and HFD neurons in the globus pallidus with 72-100% accuracy. Our results bring attention to the inherent limitations of single neuron recordings. While valuable, these recordings provide a narrow view of the neural landscape, potentially overlooking the rich tapestry of interconnected neural activity. Future investigations could benefit from simultaneous recordings of multiple neurons, which would allow for the capture of a more dynamic range of waveform metrics and better account for variables such as probe or neuron drift, which can significantly influence spike waveform and amplitude.

The adoption of multichannel approaches, such as microelectrode arrays or neuropixels[20], would allow a larger yield of simultaneous neuronal recordings. Neuron proximity to the probe can be estimated, and hence allow correction for the difference in signal quality which varies with the distance between the neuron and the probe. This would create the opportunity to use waveform and amplitude related features. Additionally, simultaneous recordings enable more insight on population activity rather than single neuron activity. .

Our findings reveal intriguing patterns surrounding lateralization: when we train models on the globus pallidus interna (GPi) as a whole, the performance is inferior compared to training on data from the left and right hemispheres separately. This suggests a potential lateralization in the GPi, where the two hemispheres exhibit distinct characteristics. In contrast, models trained on the subthalamic nucleus (STN) data as a whole outperform those trained on separate hemispheres, implying a smaller degree or lack of lateralization in the STN. Lateralization in the STN has been previously described [21,22] but structural lateralization or shifts in location is possible with age[23] which could affect the targeting accuracy and hence quality of the data, explaining why we failed to observe significant STN lateralization. In support of our findings, lateralization in GPi connectivity has been implicated in PD[24] but remains controversial[25,26].

There are two main limitations in our study. Firstly, the small dataset - approximately 16 seconds of recordings of 45 neurons per hemisphere per neuron type - limits the nature of

models we can utilize. Deep learning approaches have shown great success in neuron classification from electrophysiological recordings[27,28] but require significantly larger datasets. Additionally, the scarcity of oscillations in higher frequency bands within our dataset, particularly beta, alpha, and gamma frequencies, was puzzling. Altered beta neuronal oscillations and burst dynamics have been reported in PD[29,30]. The lack of these characteristics in our dataset could be attributed to the small sample size.

Secondly, our analysis is constrained to spike train data. Without spatial information on proximity to the source neuron, we are unable to utilize waveform and amplitude related features, as discussed above. All features extracted are hence derived solely from these spike trains. Inevitably, some features demonstrated notable interdependence (Fig. 2A, D, scatterplots). The presence of correlated features poses challenges for models like SVMs and logistic regressions, which assume feature independence, consequently diminishing their efficacy. While random forests are less affected, correlated features add less value per feature[31].

In conclusion, PD significantly impacts individuals' lives, and DBS has become a crucial therapeutic approach for managing its motor symptoms. Precise electrode implantation, particularly in target sites like the STN or GPi, is essential for optimizing therapeutic outcomes and minimizing side effects. The integration of imaging and microelectrode recordings is key to ensuring accurate targeting during the DBS procedure. While automated procedures show promise in enhancing precision, ongoing research is essential to refine techniques and further improve the efficacy of DBS in the treatment of PD.

References

1. Parkinson's Disease. *UCB* https://www.ucb-canada.ca/en/Patients/Conditions/Parkinson-s-Disease.
2. Armstrong, M. J. & Okun, M. S. Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* **323**, 548–560 (2020).
3. Poewe, W. et al. Parkinson disease. *Nature Reviews Disease Primers* **3**, 1–21 (2017).
4. Lozano, A. M., Hutchison, W. D. & Kalia, S. K. What Have We Learned About Movement Disorders from Functional Neurosurgery? *Annu. Rev. Neurosci.* **40**, 453–477 (2017).
5. Weaver, F. M. *et al.* Bilateral deep brain stimulation vs best medical therapy for patients with advanced Parkinson disease: a randomized controlled trial. *JAMA* **301**, 63–73 (2009).
6. Kalia, S. K., Sankar, T. & Lozano, A. M. Deep brain stimulation for Parkinson's disease and other movement disorders. *Curr. Opin. Neurol.* **26**, 374–380 (2013).
7. McGregor, M. M. & Nelson, A. B. Circuit Mechanisms of Parkinson's Disease. *Neuron* **101**, 1042–1056 (2019).
8. Hutchison, W. D. *et al.* Neurophysiological identification of the subthalamic nucleus in surgery for Parkinson's disease. *Ann. Neurol.* **44**, 622–628 (1998).
9. Okun, M. S. *et al.* Cognition and mood in Parkinson's disease in subthalamic nucleus versus globus pallidus interna deep brain stimulation: the COMPARE trial. *Ann. Neurol.* **65**, 586–595 (2009).
10. Hutchison, W. D. *et al.* Differential neuronal activity in segments of globus pallidus in Parkinson's disease patients. *Neuroreport* **5**, 1533–1537 (1994).
11. Wichmann, T., & DeLong, M. R. Deep brain stimulation for neurologic and neuropsychiatric disorders. *Neuron* **69**, 674–691 (2011).

12. Wichmann T, D. M. R. Deep brain stimulation for movement disorders of basal ganglia origin: restoring function or functionality? *Neurotherapeutics* **13**, 264–283 (2016).
13. Butson, C. R., Cooper, S. E., Henderson, J. M. & McIntyre, C. C. Patient-specific analysis of the volume of tissue activated during deep brain stimulation. *Neuroimage* **34**, 661–670 (2007).
14. Lozano, A. M. & Lipsman, N. Probing and regulating dysfunctional circuits using deep brain stimulation. *Neuron* **77**, 406–424 (2013).
15. Contarino, M. F. *et al.* Directional steering: A novel approach to deep brain stimulation. *Neurology* **83**, 1163–1169 (2014).
16. Tang, J. K. H. *et al.* Neuronal firing rates and patterns in the globus pallidus internus of patients with cervical dystonia differ from those with Parkinson's disease. *J. Neurophysiol.* **98**, 720–729 (2007).
17. Rajdl, K., Lansky, P. & Kostal, L. Fano Factor: A Potentially Useful Information. *Front. Comput. Neurosci.* **14**, 569049 (2020).
18. Legéndy, C. R. & Salcman, M. Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. *J. Neurophysiol.* **53**, 926–939 (1985).
19. Kaneoke, Y. & Vitek, J. L. Burst and oscillation as disparate neuronal properties. *J. Neurosci. Methods* **68**, 211–223 (1996).
20. Steinmetz, N. A. *et al.* Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science* **372**, (2021).
21. Identifying and characterizing projections from the subthalamic nucleus to the cerebellum in humans. *Neuroimage* **210**, 116573 (2020).
22. Schulz, G. M. *et al.* Selective left, right and bilateral stimulation of subthalamic nuclei in Parkinson's disease: differential effects on motor, speech and language function. *J. Parkinsons. Dis.* **2**, 29–40 (2012).

23. Pereira, J. L. B. *et al.* Lateralization of the Subthalamic Nucleus with Age in Parkinson's Disease. *Basal Ganglia* **6**, 83–88 (2016).

24. Miranda-Domínguez, Ó. *et al.* Lateralized Connectivity between Globus Pallidus and Motor Cortex is Associated with Freezing of Gait in Parkinson's Disease. *Neuroscience* **443**, 44–58 (2020).

25. Lin, Z., Zhang, C., Li, D. & Sun, B. Lateralized effects of deep brain stimulation in Parkinson's disease: evidence and controversies. *npj Parkinson's Disease* **7**, 1–8 (2021).

26. Sharim, J., Yazdi, D., Baohan, A., Behnke, E. & Pouratian, N. Modeling Laterality of the Globus Pallidus Internus in Patients With Parkinson's Disease. *Neuromodulation* **20**, 238–242 (2017).

27. Yip, M. C., Gonzalez, M. M., Valenta, C. R., Rowan, M. J. M. & Forest, C. R. Deep learning-based real-time detection of neurons in brain slices for in vitro physiology. *Sci. Rep.* **11**, 1–9 (2021).

28. Ophir, O., Shefi, O. & Lindenbaum, O. Neuronal Cell Type Classification using Deep Learning. *arXiv [cs.LG]* (2023).

29. Yu, Y. *et al.* Parkinsonism Alters Beta Burst Dynamics across the Basal Ganglia-Motor Cortical Network. *J. Neurosci.* **41**, 2274–2286 (2021).

30. Asadi, A., Madadi Asl, M., Vahabie, A.-H. & Valizadeh, A. The Origin of Abnormal Beta Oscillations in the Parkinsonian Corticobasal Ganglia Circuits. *Parkinsons Dis.* **2022**, 7524066 (2022).

31. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).

Appendix

Table 1: List of features and description

	Feature	**Description**
1	ifr_mean	Mean of instantaneous firing rate
2	ifr_skew	Skew of instantaneous firing rate distribution
3	ifr_kurtosis	Kurtosis of instantaneous firing rate distribution
4	ifr_fano_factor	Fano factor of instantaneous firing rate
5	ISI_mean	Mean of inter spike intervals (ISI)
6	ISI_skew	Skew of ISI distribution
7	ISI_kurtosis	Kurtosis of ISI distribution
8	ISI_cv	Coefficient of variance of ISI distribution
9	num_bursts	Number of burst events (minimum of 3 spikes) as detected by poisson surprise method
10	mean_surprise	Mean surprise value of bursts
11	burst_index	Ratio of mean ISI to median ISI
12	mean_burst_duration	Mean length of bursts
13	var_burst_duration	Variance of length of bursts distribution
14	mean_interburst_duration	Mean of time between bursts
15	var_interburst_duration	Variance of time between bursts distribution
16	max_peak_freq	The highest frequency detected in the autocorrelogram or periodogram of a spike train, indicating the dominant oscillatory activity in the neuronal pattern.
17	delta_band	Presence of delta band oscillations
18	theta_band	Presence of theta band oscillations
19	alpha_band	Presence of alpha band oscillations
20	beta_band	Presence of beta band oscillations
21	gamma_band	Presence of gamma band oscillations

Table 2: Test set accuracy results (highest test set performance in bold)

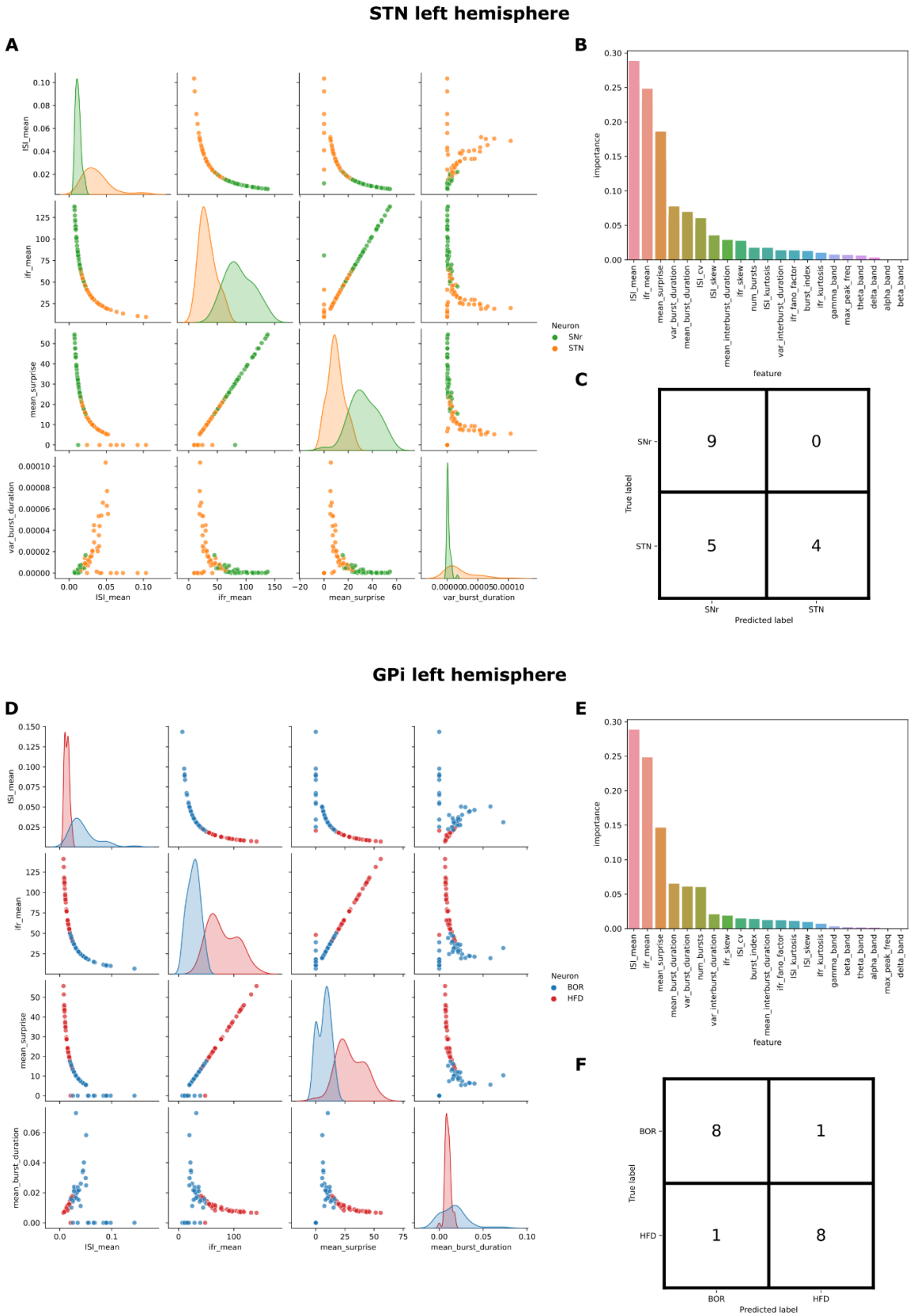
	Logistic regression		Random forest		SVM	
	Train	Test	Train	Test	Train	Test
STN Left	94	**72**	96	**72**	94	**72**
STN Right	96	89	100	**94**	94	89
GPi Left	100	78	100	**89**	97	83
GPi Right	97	94	96	**100**	97	89

Table 3: Model trained on data pooled across hemispheres

	Random forest	
	Train accuracy	Test accuracy
STN	98	95
GPI	98	89

Table 4: Model trained and tested on different hemispheres

Trained on	**Tested on**	**Random forest accuracy**
STN left	STN right	89
STN right	STN left	91
GPi left	GPi right	97
GPi right	GPi left	92



Supplementary Figure 1: Classification of Neuron type in subthalamic nucleus and globus pallidus of the left hemisphere. (A, D) Selected features show separation between

neuron types. **(B, E)** Features ranked by Gini importance. **(C, F)** Confusion matrix of test set results.