Micro-Electrode Recordings – STN-DBS Surgery

Classification of Micro-Electrode Recordings for optimizing Subthalamic Nucleus Targeting

Group #14: Debellis Angela, Lionetti Irene, Petruzzella Rossella, Vitulano Manuela

***Abstract*— Deep Brain Stimulation (DBS) of the subthalamic nucleus (STN) is a widely adopted surgical intervention for managing motor symptoms in Parkinson's disease (PD), including tremor, rigidity, and bradykinesia. Precise localization of the STN during electrode implantation is critical to the success of the procedure, yet it remains challenging due to the complexity of distinguishing STN signals from those of adjacent brain areas. Microelectrode recordings (MER) obtained during surgery exhibit unique signal patterns that can assist in identifying the STN. This study evaluates and compares the performances of the k-Nearest Neighbor (kNN) classifier and Bayesian classifier in accurately detecting the STN region using MER data. A dataset comprising MER signals from 3 PD patients was preprocessed, and 11 key features, spike dependent and independent, were extracted to train and validate the two models. Performance analysis confirmed the superiority of the kNN classifier. These findings demonstrate the potential of kNN as a robust tool for aiding neurosurgeons in DBS procedures, ensuring precise STN targeting while reducing uncertainty in electrode placement and timing of the procedure.**

***Index Terms*— Parkinson's Disease, MERs, STN-DBS Surgery, Classifiers, Targeting**

I. Introduction

PARKINSON'S disease (PD) is a common neurodegenerative disorder characterized by the degeneration of the pars compacta of the substantia nigra and the striatal pathway. This leads to increased inhibitory projections from the basal ganglia nuclei to the cortex and brainstem, causing abnormal activation of dopaminergic and non-dopaminergic structures involved in movement. Key symptoms include bradykinesia, rigidity, tremor, postural instability, cognitive decline, and depression. Among these, gait disturbances—such as slow and reduced locomotion, diminished stride length, lower velocity despite normal cadence, and increased risk of falls—are major contributors to disability as the disease progresses.

High-frequency bilateral deep brain stimulation (DBS) of the subthalamic nucleus (STN) has emerged as a well-established surgical therapy for managing advanced PD, particularly for patients whose motor complications cannot be adequately controlled with medication. DBS has been demonstrated to effectively mitigate all cardinal symptoms of PD—rigidity, bradykinesia, and tremor—as well as drug-induced motor complications such as fluctuations and dyskinesias [1]. Precise localization of the STN is critical for successful DBS outcomes. Accurate targeting reduces the duration of the procedure, minimizes the risk of adverse effects, and improves clinical outcomes. However, this task is challenging due to the complexity of distinguishing STN signals from those of adjacent brain regions. Neurosurgeons rely on a combination of preoperative imaging, intraoperative tomography, and Micro-Electrode Recordings (MER) to identify the STN. Among these, MER signals, which reflect the unique neural activity of different brain regions, have proven particularly valuable. Each structure traversed during electrode implantation—such as the anterior thalamus, zona incerta, and substantia nigra—exhibits distinct signal patterns that can aid in identifying the STN [2], [3].

Supervised machine learning algorithms offer a promising solution for automating and enhancing the accuracy of STN localization using MER data. By analyzing characteristic features extracted from MER signals, these algorithms can classify data as originating from within or outside the STN. This study investigates the performance of two classifiers—k-Nearest Neighbors (kNN) and Bayesian classifiers—in distinguishing STN signals. By employing a dataset of MER signals from patients with advanced PD, we aim to assess the classifiers' accuracy and robustness in aiding neurosurgeons during DBS procedures. Our findings highlight the potential of these methods to optimize electrode placement, reduce surgical uncertainty, and ultimately improve patient outcomes [4], [5].

II. Materials and Methods

A. Data Acquisitions

MERs were collected from three patients with advanced Parkinson's Disease (PD) undergoing bilateral DBS of the STN. The data, anonymized for patient privacy, were obtained using the NeuroSet system (Neurostar, Tübingen, Germany) at a sampling frequency of 20 kHz. Signals were recorded from both hemispheres, at depths ranging from -8 mm to +5 mm relative to the surgical zero-point, identified via preoperative MRI. Recording durations varied between a few seconds and one minute, depending on the depth. To ensure compatibility for analysis, all signals were high-pass filtered at 200 Hz and converted to microvolts (μV), effectively removing low-frequency noise without compromising the integrity of the neural activity data.

B. Data Organization

The dataset for this study comprised three MATLAB® structures (called *"Data.mat"*) corresponding to patients D20, D33, and D38. Each structure included the following fields:

• *Sampl_Freq*: The sampling frequency of the recordings.
• *LeftHemisphere* and *RightHemisphere*: Data from the left and right hemispheres, respectively.
For each hemisphere, the following subfields were defined:
• *Num_MERs*: Total number of signals acquired.
• *MERs*: A collection of recording vectors, each labeled by its relative position from the surgical zero-point. The labels use a format of *mx* or *px*, where m (minus) and p (plus) denote the depth relative to the zero-point in millimeters. For example, *m3* refers to a recording obtained at -3 mm;
• *Target*: A binary vector indicating the classification of each signal. A value of 1 represents signals acquired within the STN, while 0 corresponds to signals from outside the STN. The length of this vector matches the total number of signals recorded (*Num_MERs*).

C. Filtering and Segmentation

Preprocessing involved applying an Infinite Impulse Response (IIR) band-pass filter with a frequency range of 200–6000 Hz, chosen based on the characteristic frequency band of the intracortical signal as observed through the power spectral density (PSD) analysis. This filtering step effectively removed low-frequency noise and high-frequency artifacts from the raw signals, ensuring that essential signal features were preserved while reducing interference [6].

Filtered signals were visually examined to validate the noise reduction process. For example, as can be seen in Fig.1, recordings obtained within the STN exhibited distinct characteristics, such as higher amplitude (peaks in a range between [-15,+20] μV), lower variability, and elevated non-linear energy, indicative of increased neuronal synchronization typical of Parkinson's disease. In contrast, signals recorded outside the STN, in Fig.2, displayed lower amplitude (peaks around [-15,+15] μV), reduced spike density, and diminished energy, reflecting less synchronized neural dynamics. These features are critical for accurately identifying the STN during DBS.

To prepare the data for analysis and improve classification robustness, the filtered signals were segmented into overlapping 1-second epochs with a 50% overlap. This approach augmented the dataset and standardized input lengths for feature extraction. The number of resulting epochs varied according to the original signal duration, ensuring comprehensive coverage of the recordings.

D. Features Extraction

For each epoch, the following features were extracted [3], [4], [7]:
1. Spike-independent features:

- Kurt: Signal kurtosis.
- CL: Sum of absolute value for the signal derivative.
- TH: Threshold in the signal amplitude.
- PK: Peaks number.
- RA: RMS value of amplitude.
- NE Non-linear mean energy.

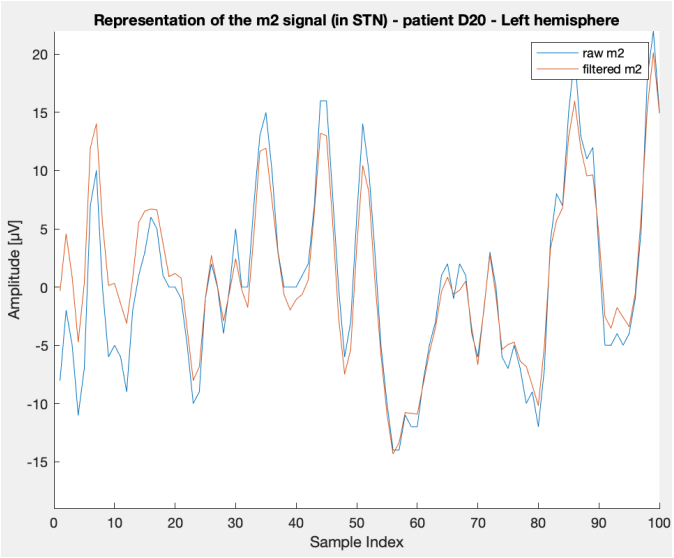

Fig. 1. Raw and Filtered MER from patient D20 inside STN region

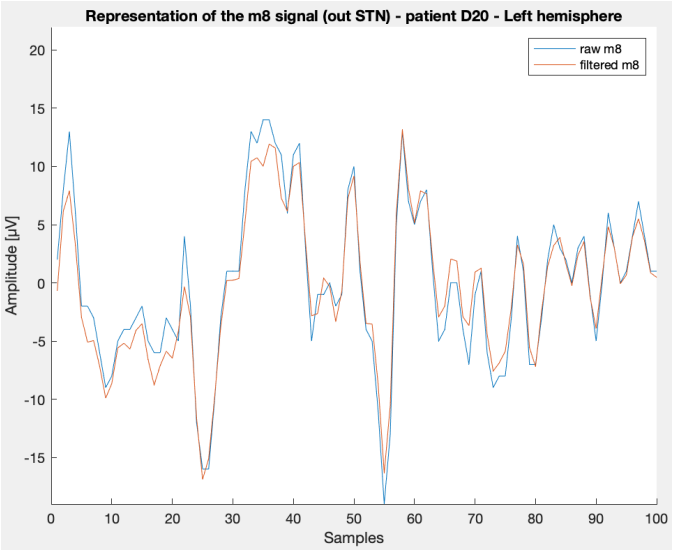

Fig. 2. Raw and Filtered MER signal from patient D20 outside STN region

- ZC: Zero crossings.

2. Spike-dependent features:

- SC: Spikes per second.
- SMAD: Mean spike differential amplitude as the mean value for the amplitude in two consecutive spikes.
- SSD: Standard deviation of interspike intervals.
- SF: Mean spike trigger frequency.

E. Classifiers

After the feature extraction process, the resulting data, from every hemisphere of each patient, were organized into a matrix, where each row represented an epoch of the filtered signals, and each column corresponded to one of the 11 extracted features. An additional column was included to label each epoch as either 0 or 1, according to the corresponding

value in the target vector, indicating whether the signal originated inside or outside the STN.

To evaluate model performance, the dataset was partitioned using a leave-one-patient-out cross-validation approach, a specialized variation of the leave-one-out method particularly suited for medical data. This methodology ensures that all data from a single patient is excluded from the training process and reserved exclusively for testing, thus preventing information leakage and providing a robust evaluation of the model's generalization capabilities.

Three experimental scenarios were conducted, with the following test sets:

1) Patient D20 served as the test set
2) Patient D33 served as the test set
3) Patient D38 served as the test set

For each scenario, the training and validation sets were constructed from the data of the remaining two patients. The training set consisted of 60 % of the epochs from the underrepresented class (class 0) and an equal number of epochs randomly selected from the majority class (class 1) to address potential class imbalance. The validation set included the remaining epochs, ensuring independence from the training set.

Two classifiers were employed for signal classification: the kNN algorithm and the Naive Bayes classifier.

The kNN algorithm is a non-parametric, supervised learning method that classifies data based on the majority class among the nearest k neighbors in the feature space. For this study, the optimal k was determined through the equation (1), where n represents the total number of epochs in the training set:

$$k = \sqrt{n} \tag{1}$$

In contrast, the Naive Bayes classifier utilizes a probabilistic framework based on Bayes' theorem (2). This method assumes conditional independence among features given the class label, which significantly simplifies the computation.

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{2}$$

By applying these two methods, the study aimed to determine which classifier demonstrated superior accuracy and robustness in distinguishing STN signals from non-STN signals. This comparison provided insights into their applicability in assisting neurosurgeons during DBS procedures.

III. Results

To evaluate the performances of each classifier for every experimental scenario and every subset, confusion matrices were produced. In the following, results obtained with the two classifiers were compared by calculating relevant performance metrics, which were extracted from the confusion matrices.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1_{S}core = \frac{2 * TP}{2 * TP + FP + FN} \tag{5}$$

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \tag{6}$$

$$ErrorRate = \frac{FP + FN}{TN + TP + FP + FN} \tag{7}$$

$$FPR = \frac{FP}{FP + TN} \tag{8}$$

$$FNR = \frac{FN}{TP + FN} \tag{9}$$

Considering the three experimental setups using the leave-one-out method, the final performance was determined by calculating the mean and standard deviation of the mentioned metrics for the training, validation, and test sets. The results are summarized in Tables I, II, and III.

To visualize the performance differences between the two classifiers across these sets, histogram representations are provided in Fig.3, Fig.4, and Fig.5.

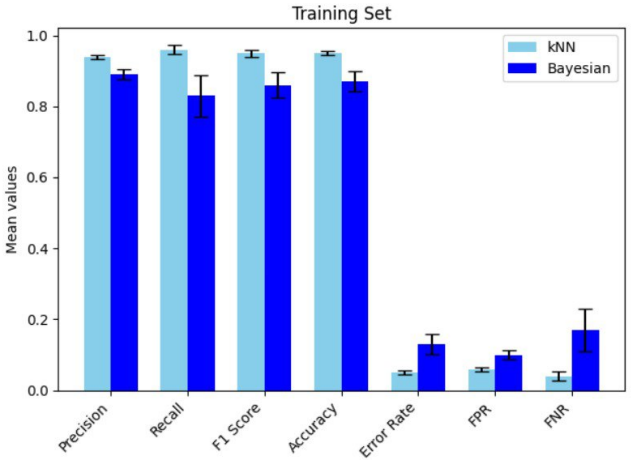

Fig. 3. Histogram of performances of Training Set.

TABLE I
MEAN AND STANDARD DEVIATION VALUES FOR TRAINING SET RESULTS.

Training Set	Precision	Recall	F1 Score	Accuracy	Error Rate	FPR	FNR
kNN	0.94 ± 0.0058	0.96 ± 0.012	0.95 ± 0.010	0.95 ± 0.0058	0.05 ± 0.0058	0.06 ± 0.0058	0.04 ± 0.012
Bayesian	0.89 ± 0.015	0.83 ± 0.059	0.86 ± 0.035	0.87 ± 0.029	0.13 ± 0.029	0.10 ± 0.012	0.17 ± 0.059

TABLE II
MEAN AND STANDARD DEVIATION VALUES FOR VALIDATION SET RESULTS.

Validation Set	Precision	Recall	F1 Score	Accuracy	Error Rate	FPR	FNR
kNN	0.81 ± 0.070	0.97 ± 0.032	0.88 ± 0.027	0.94 ± 0.010	0.06 ± 0.010	0.067 ± 0.0058	0.033 ± 0.032
Bayesian	0.71 ± 0.10	0.83 ± 0.046	0.77 ± 0.072	0.89 ± 0.0058	0.11 ± 0.0058	0.10 ± 0.015	0.17 ± 0.046

TABLE III
MEAN AND STANDARD DEVIATION VALUES FOR TEST SET RESULTS.

Test Set	Precision	Recall	F1 Score	Accuracy	Error Rate	FPR	FNR
kNN	0.87 ± 0.085	0.86 ± 0.19	0.85 ± 0.076	0.89 ± 0.057	0.11 ± 0.057	0.14 ± 0.17	0.13 ± 0.20
Bayesian	0.83 ± 0.070	0.76 ± 0.30	0.77 ± 0.072	0.86 ± 0.031	0.14 ± 0.031	0.16 ± 0.17	0.25 ± 0.30

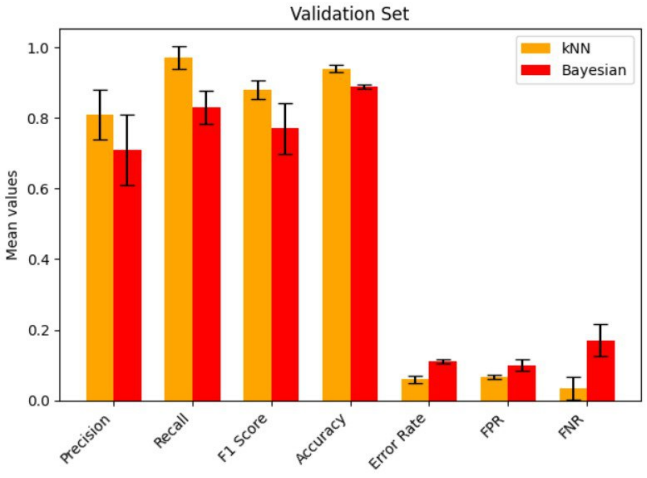

Fig. 4. Histogram of performances of Validation Set.

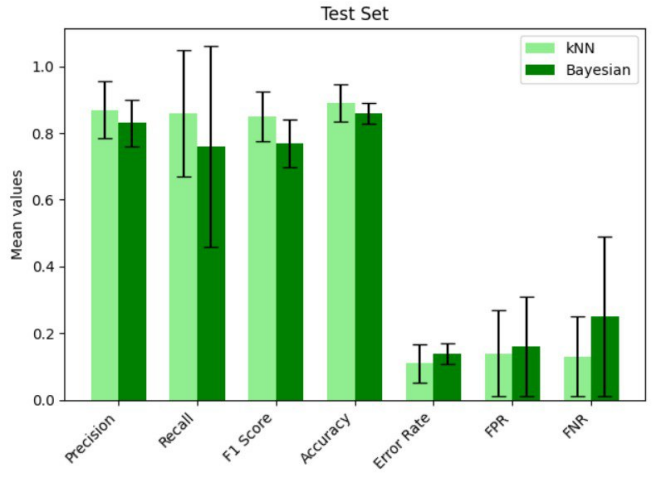

Fig. 5. Histogram of performances of Test Set.

Following, there is an analysis of the results for the three subsets:

Training Set: The kNN classifier outperformed the Bayesian classifier across all metrics. Specifically, it showed higher Precision, Recall, F1 Score, and Accuracy, while demonstrating lower Error Rate, FPR, and FNR compared to the Bayesian classifier.

Validation Set: The kNN continued to outperform the Bayesian classifier, particularly in Precision, F1 Score, and Accuracy. However, the performance gap between the two classifiers decreased compared to the training set.

Test Set: The kNN classifier again showed better performance in all metrics, although the differences between the classifiers were less significant, especially in terms of Recall and Precision.

The difference between the Mean-Square-Error (MSE) (10) of the training set and the test set was computed for each of the three trials, and the RMS (11) was computed too, with the results shown in the table (Table IV). A negative value indicates that the MSE is generally lower in the training set compared to the test set, except for the Bayesian classifier when applied to the dataset of patient D33 as the test set.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{10}$$

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}x_i^2} \tag{11}$$

IV. DISCUSSION AND CONCLUSION

This study aimed to optimize the targeting of the STN using MER data during DBS surgery for PD. The performance of two classifiers, kNN and Bayesian, was compared to assess their effectiveness in accurately identifying the STN and distinguishing it from adjacent brain regions. Based on the analysis of MSE differences between the training and test sets, along with primary metrics (Precision, Recall, F1 Score, and Error Rate), several key observations emerged.

The kNN classifier showed consistently higher MSE differences between training and test sets, suggesting slight instability and a tendency for overfitting. However, it achieved higher Precision, Recall, and F1 Scores compared to the Bayesian classifier, with relatively low standard deviations across training, validation, and test sets. This indicates that the kNN model exhibits more repeatability.

In contrast, the Bayesian classifier demonstrated lower MSE differences, suggesting a more stable generalization between datasets. It showed lower RMS values, indicating less fluctuation in error between training and test sets. While its overall performance metrics were slightly lower, the Bayesian model's ability to maintain consistent predictions, particularly with its lower Error Rate and higher stability, made it a strong contender for generalizing across different datasets.

TABLE IV
DIFFERENCE BETWEEN TRAINING ERROR AND TEST ERROR

Classifiers	Setup 1	Setup 2	Setup 3	RMS
kNN	-0.030	-0.040	-0.11	0.070
Bayesian	-0.040	0.060	-0.030	0.045

Overall, both classifiers exhibit strengths, but the kNN classifier is more effective in achieving higher accuracy in identifying the STN, while the Bayesian classifier excels in generalization and stability.

In conclusion, the key takeaway from this study is that optimizing the targeting of the STN for DBS surgery requires not only high-performance classifiers but also careful attention to data pre-processing to ensure optimal classifier performance.

An interesting direction for future research is the development of real-time classification methods. Classifiers that don't rely on extensive pre-processing could bring significant benefits, such as faster processing times and the ability to quickly adapt to changes in data during surgery. A real-time system could also simplify the procedure by reducing the need for pre-surgical data preparation, making it more efficient and less time-consuming. While this approach might come with a slight trade-off in accuracy compared to pre-processed data, the gains in operational efficiency and adaptability to real-time changes in MER data make it a promising area worth exploring.

Moreover, future studies should aim to combine the strengths of both classifiers to balance accuracy and generalization, improving the outcomes of DBS procedures for treating Parkinson's disease.

V. REFERENCES

[1] M. G. et al., "Muscle synergies in parkinson's disease before and after the deep brain stimulation of the bilateral subthalamic nucleus," *Scientific Reports*, 2023.

[2] J. M. B. et al., "Deep brain stimulation for parkinson disease: an expert consensus and review of key issues," *Arch. Neurol.*, vol. 68, p. 165, 2011.

[3] J. R.-M. A. Guerrero-Martínez, "Characterization of microelectrode records in deep brain stimulation applied to parkinson's disease patients," *IFMBE Proceedings*, vol. 49, p. 647–650, 2015.

[4] L. S. et al., "Stn area detection using k-nn classifiers for mer recordings in parkinson patients during neurostimulator implant surgery," *J. Phys.: Conf. Ser.*, vol. 705, p. 012050, 2016.

[5] K. et al., "Unsupervised clustering of micro-electrophysiological signals for localization of subthalamic nucleus during dbs surgery," *9th International IEEE EMBS Conference on Neural Engineering, San Francisco, CA, USA*, 2019.

[6] R. et al., "Advancements in digital signal processing: Design and implementation of iir bandpass optical filter using optisystem," *International Conference of Advances Technologies for Health Engineering Science*, 2024.

[7] W. A. C. et al., "Pattern recognition approaches for identifying subcortical targets during deep brain stimulation surgery," *IEEE Intell. Syst.*, vol. 26, p. 54–63, 2011.