Clinical Neurophysiology 130 (2019) 145–154

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

A review on microelectrode recording selection of features for machine learning in deep brain stimulation surgery for Parkinson's disease

Kai Rui Wan [a,b], Tomasz Maszczyk [c], Angela An Qi See [a,b], Justin Dauwels [c], Nicolas Kon Kam King [a,b,d,*]

[a] Department of Neurosurgery, National Neuroscience Institute, Singapore
[b] Department of Neurosurgery, Singapore General Hospital, Singapore
[c] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore
[d] Duke-NUS Medical School, Singapore

ARTICLE INFO

Article history:
Accepted 17 September 2018
Available online 25 September 2018

Keywords:
Parkinson's disease
Machine learning
Microelectrode recording
Automation

HIGHLIGHTS

- Automatic STN detection improves the efficiency and accuracy of target location.
- Both spike independent and dependent features are useful for detection of STN.
- Further research required for STN sub territories.

ABSTRACT

Objective: This study seeks to systematically review the selection of features and algorithms for machine learning and automation in deep brain stimulation surgery (DBS) for Parkinson's disease. This will assist in consolidating current knowledge and accuracy levels to allow greater understanding and research to be performed in automating this process, which could lead to improved clinical outcomes.
Methods: A systematic literature review search was conducted for all studies that utilized machine learning and DBS in Parkinson's disease.
Results: Ten studies were identified from 2006 utilizing machine learning in DBS surgery for Parkinson's disease. Different combinations of both spike independent and spike dependent features have been utilized with different machine learning algorithms to attempt to delineate the subthalamic nucleus (STN) and its surrounding structures.
Conclusion: The state-of-the-art algorithms achieve good accuracy and error rates with relatively short computing time, however, the currently achievable accuracy is not sufficiently robust enough for clinical practice. Moreover, further research is required for identifying subterritories of the STN.
Significance: This is a comprehensive summary of current machine learning algorithms that discriminate the STN and its adjacent structures for DBS surgery in Parkinson's disease.

1. Introduction

Parkinson's disease (PD) is a progressive, neurodegenerative disease characterized by motor symptoms that include bradykinesia, resting tremor, postural instability, and rigidity leading to significant effects on patient's quality of life. (Brocker et al., 2017; Gulberti et al., 2015). High frequency deep brain stimulation (DBS), to either the globus pallidus internus (GPi) or subthalamic nucleus (STN) (Williams et al., 2014), is an effective surgical treatment for moderate to advanced PD that improves motor symptoms and quality of life (Brocker et al., 2017; Holt and Netoff, 2016; Karamintziou et al., 2016, 2017; Kühn and Volkmann, 2017; Rowland et al., 2017).

The STN can be heuristically divided into its sensorimotor, associative and limbic regions (Lourens et al., 2013) and stimulation at different STN subterritories will result in differential effects on motor and emotional functions (Mallet et al., 2007; Pozzi et al., 2016). Sub-optimal positioning of DBS electrodes account for up to 40% of cases of inadequate postoperative stimulation efficacy (Okun et al., 2005) and other associated limbic or cognitive side effects (Holt and Netoff, 2016; Shamir et al., 2012; Telkes et al.,

* Corresponding author at: Consultant Neurosurgeon, Department of Neurosurgery, National Neuroscience Institute, 11 Jalan Tan Tock Seng, Singapore 308433, Singapore.
E-mail address: nicolas.kon.k.k@singhealth.com.sg (N.K.K. King).

https://doi.org/10.1016/j.clinph.2018.09.018
1388-2457/

2014; Teplitzky et al., 2016; Valsky et al., 2017; van Wijk et al., 2017). The efficacy of DBS is thus dependent on careful selection of the appropriate patient (Rowland et al., 2017), as well as accurate localization of this extremely small target nucleus (Cagnan et al., 2011; Taghva, 2011; Telkes et al., 2014; Teplitzky et al., 2016; van Wijk et al., 2017; Zaidel et al., 2009).

The current practice for target localization in DBS involves a multi-modal approach, which typically includes the use of stereotactic atlas coordinates, brain imaging, intraoperative microelectrode recording (MER) and clinical assessment of the electrode placement with intraoperative electrical stimulation (Lima, 2015; Snellings et al., 2009; Telkes et al., 2014). Real-time testing and accurate physiological delineation of the functional boundaries of the STN and its surrounding structures (Novak et al., 2011) are two key advantages of MER that can reduce targeting errors by firstly accounting for differences in neuroanatomy (Lima, 2015); secondly, overcoming resolution limitations of neuroimaging (Chaovalitwongse et al., 2011) and thirdly, compensate for anatomic shifts during surgery due to CSF leak, pneumocephalus (D'Haese et al., 2012) and brain deformation by tissue compression (Cagnan et al., 2011; Chaovalitwongse et al., 2011; Guo et al., 2007; Rajpurohit et al., 2015).

The MER signals are monitored and assessed intraoperatively by either the neurosurgeon and/or neurophysiologist (Cagnan et al., 2011; Falkenberg et al., 2006a,b; Snellings et al., 2009; Telkes et al., 2014). The STN region is identified by an abrupt increase in firing rates, irregular bursting activity, background neural activity and characteristic spike patterns (Falkenberg et al., 2006a,b; He, 2009; Lima, 2015; Moran et al., 2006) (see Fig. 1). Kinesthetic activity, which is induced by the passive movement of the contralateral limbs, causes alteration of neural firing (Novak et al., 2011) which can assist in establishing the location of the STN.

Nevertheless, there are several challenges with the usage of MER. These subjective and experience-based techniques (Cagnan et al., 2011; Snellings et al., 2009; Telkes et al., 2014) possess inherent limitations (Snellings et al., 2009) with possible inconsistencies in its analysis due to its critical reliance on user expertise and judgment. The sheer volume and complexity of data can be perceived intraoperatively in a single time domain only (Lima, 2015) and other anatomic challenges such as an uninterrupted subthalamic nucleus to substantia nigra transition and gaps within the STN (Moran et al., 2006) could lead to erroneous labeling of the structures (Valsky et al., 2017). Furthermore, complex signal patterns, noise interference or drift as well as artifacts from patient movement, blood or edema may also exponentially increase the difficulty of MER data interpretation (Telkes et al., 2014; Valsky et al., 2017).

In recent years, there has been much interest in the automation of MER-based intraoperative targeting during DBS surgery (Falkenberg et al., 2006b; Moran et al., 2006; Rajpurohit et al., 2015; Telkes et al., 2014). An online real-time implementation will allow input data to be processed within a short frame of time to be available as feedback to the algorithm for analysis. The potential benefits of online real-time automation in MER-based targeting include: reducing surgical time (Cardona et al., 2012), reducing dependence on neurophysiological expertise and providing consistent and reliable identification of a well-defined quantifiable target to place the DBS electrode (Falkenberg et al., 2006b; Snellings et al., 2009; Telkes et al., 2014). More importantly, automation may also assist in decreasing the patient's burden, who are awake and usually in the 'off state' during the surgery (Falkenberg et al., 2006b; He, 2009).

The aims of this review are to critically appraise the available literature on automating the MER process in DBS surgery for PD patients as well as identify potential areas for further research.

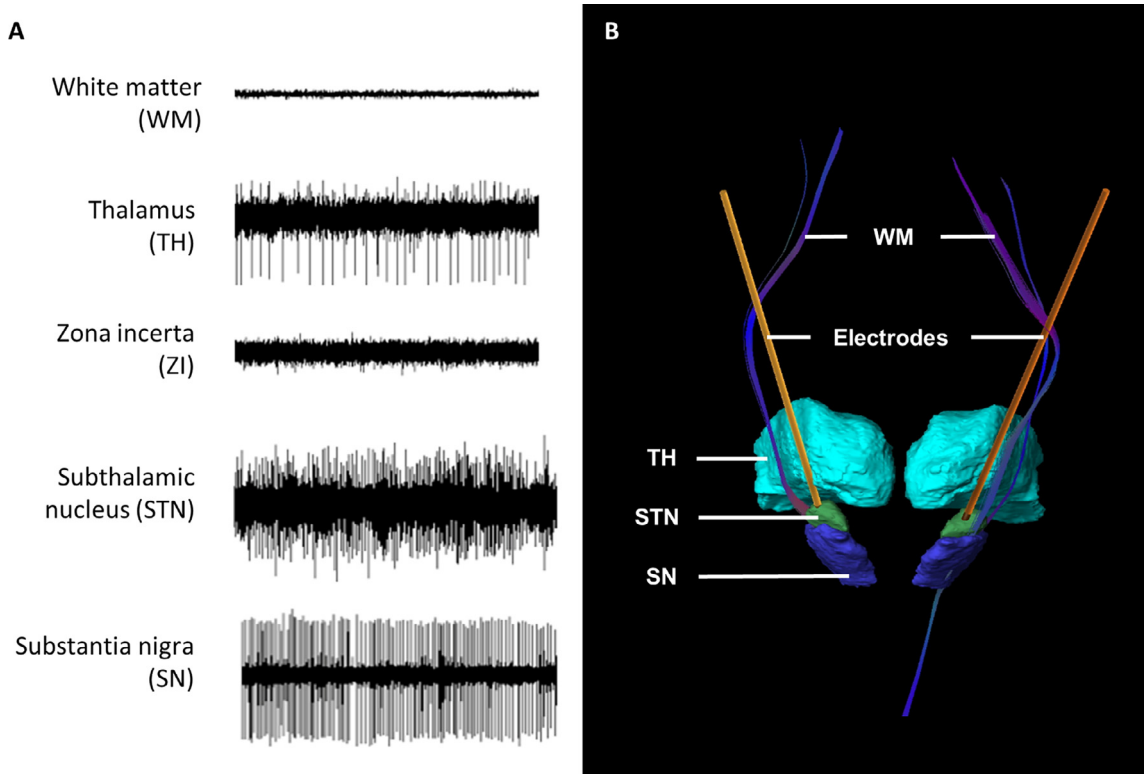

Fig. 1. 3D image of STN (left) with MER recordings (right) in a typical trajectory. (A) Upon entering the subthalamic nucleus, there is an increased compound firing rate and neural noise background, which decreases upon exiting the STN. (B) The typical DBS trajectory will traverse the thalamus (TH), followed by zona incerta (ZI), arriving at the subthalamic nucleus (STN), exiting into white matter (WM), and ending at the substantia nigra (SN).

2. Overview of machine learning and classification algorithms

Machine learning (ML) is a branch of artificial intelligence that enables computer algorithms to learn from experience without explicitly being programmed (Cabitza and Banfi, 2017; Senders et al., 2017). The field of ML is broadly divided into supervised and unsupervised learning.

Supervised learning algorithms learn from "labeled" training data to produce a model that can make predictions on previously unseen data. The desired output of these training data is known. In the learning process, algorithms attempt to find the optimal combination of input variables (also known as features) and weights given to these features in order to perform accurate predictions (Senders et al., 2017). Labeling can be used to train the classifier and test its estimation performance. However, it may also be an additional source of additional (human) error, bias or misconception, which could obscure the ground truth (Lima, 2015). In this respect, classification based on manual annotations is only as good as the human evaluation itself (Lima, 2015). With unsupervised learning techniques, only unlabeled data are available and the algorithm then seeks to find similarities and patterns (Senders et al., 2017). Data processing might therefore help to extract meaningful, and at present hidden, information from MER data to improve current treatment modalities (Chaovalitwongse et al., 2011). Examples of the machine learning classifiers are presented in Fig. 2.

2.1. Naïve Bayes classifier

Bayesian inference is based on Bayes rule (Chaovalitwongse et al., 2011):

$$p(c|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|c)p(c)}{p(\boldsymbol{x})}.$$

Bayes rule describes how the probability of an event $\boldsymbol{x}$ changes from a prior $p(\boldsymbol{x})$, before we observe anything, to a posterior $p(\boldsymbol{x}|c)$ once we have observed event c (Moran et al., 2006). The naïve Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature, and therefore, all these properties (features) independently contribute to the probability.

2.2. Decision tree and random forest

A classification tree is a non-parametric discrimination method to recursively partition data samples into two or more groups based on a specific splitting criterion. The most common criteria are Gini Index, Information Gain, and Gain Ratio. The Gini index measures the impurity of a data partition D and is defined as:

$$Gini = 1 - \sum_{i=1}^{k} p_i^2,$$

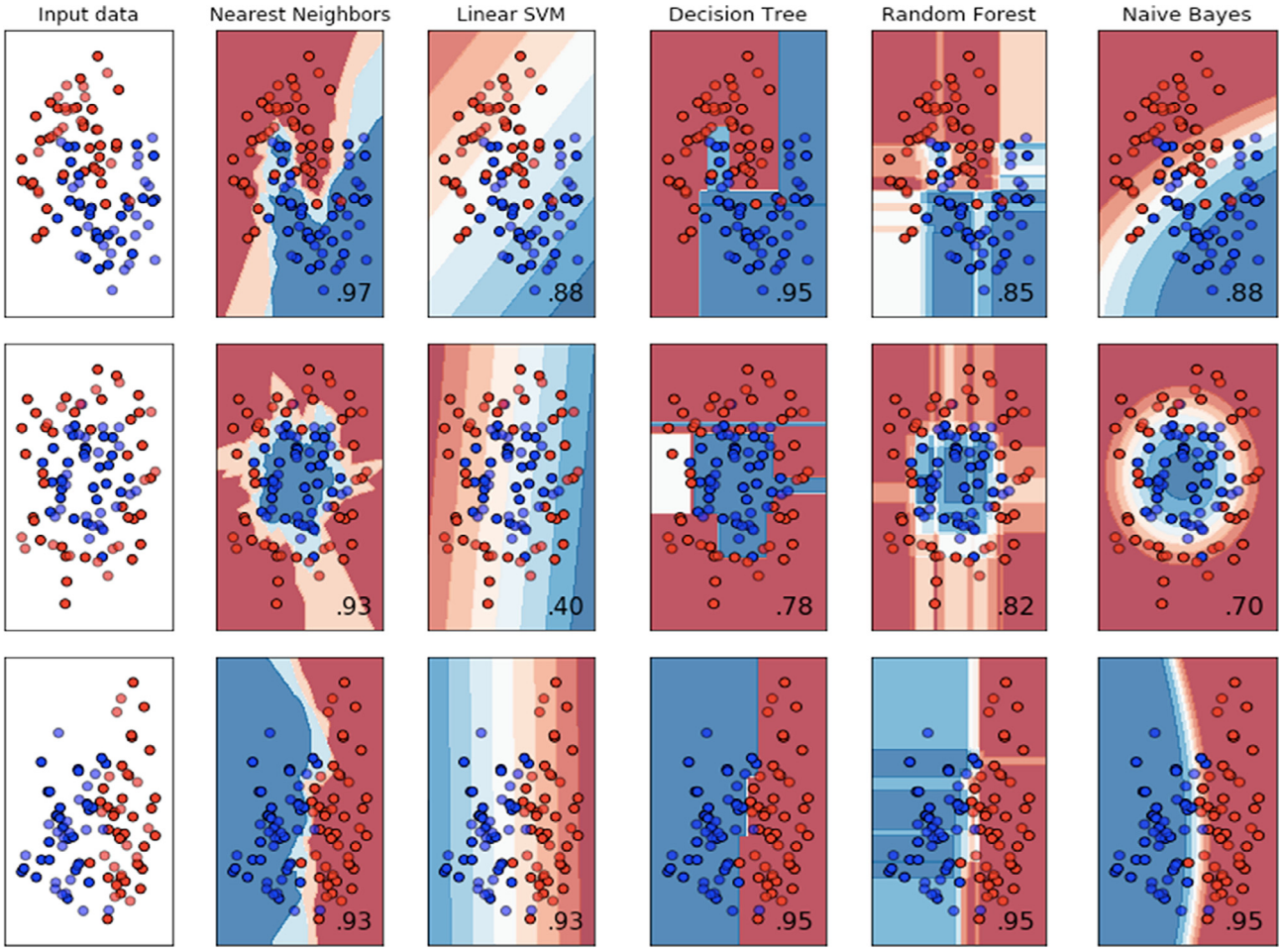

Fig. 2. Models of machine learning, adapted from an image by scikit-learn.org. (Pedregosa, 2011). Raw input data (first column) is applied to different machine learning algorithms, including nearest neighbours, support vector machine, decision trees, random forest, and naïve Bayes. The classification accuracies of the given example input data (first column) are reflected at the bottom right corner of the subsequent diagrams. The accuracy is however dependent on the type of data supplied to each machine learning algorithm, as evident by the three different examples of input data.

where k is the number of classes. The Gini index considers a binary split for each attribute, and selects for splitting the attribute that maximizes the reduction in impurity.

Information Gain is defined as:

$$\Delta H = H - \frac{m_L}{m} H_L - \frac{m_R}{m} H_R,$$

where m is the total number of instances, m_k number of instances belonging to class k, and entropy H is defined as:

$$H = -\sum_{i=1}^{k} p_k log_2 p_k.$$

Gain ratio is a modification of Information Gain defined as:

$$GainRatio = \frac{Gain}{SplitInfo},$$

where split information value represents the potential information generated by splitting the training data D into v partitions:

$$SplitInfo = -\sum_{i=1}^{v} \frac{|D_i|}{|D|} log_2\left(\frac{|D_i|}{D}\right).$$

The construction of the classification tree involves a selection of the splits, a decision on a terminal node, and class assignment to each terminal node (Chaovalitwongse et al., 2011). The topmost decision node in a tree, which corresponds to the best predictor, is called root node.

A random forest is a conglomeration of multiple classification trees.

Decision trees can handle both categorical and numerical data. The disadvantages of decision trees are the heavy memory requirements and training times. Boosted decision trees can help to avoid overfitting of data (Friedman, 1999).

2.3. k-Nearest Neighbour (kNN)

kNN is a supervised non-parametric classifier method. Unlabeled samples (test data) are labelled based on their similarity to samples in the training data (Chaovalitwongse et al., 2011). This method assigns distance functions from the k nearest neighbors to the sample.

The most common distance functions include:

- Euclidean metric: $D(\boldsymbol{x},\boldsymbol{y}) = \sqrt{\sum_{i=1}^{d}(x_i - y_i)^2}$
- Minkowski metric: $D(\boldsymbol{x},\boldsymbol{y})^{\alpha} = \sum_{i=1}^{d}|x_i - y_i|^{\alpha}$
- Mahalanobis distance: $D(\boldsymbol{x},\boldsymbol{y}) = \sqrt{(\boldsymbol{x}-\boldsymbol{y})'\boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{y})}$

where $\mathbf{C}$ is the covariance matrix

- Cosine distance: $D(\boldsymbol{x},\boldsymbol{y}) = \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{||\boldsymbol{x}||||\boldsymbol{y}||}$
- Hamming distance: $D(\boldsymbol{x},\boldsymbol{y}) = \frac{\#(x_i \neq y_i)}{d}$
- Correlation distance: $D(x,y) = \frac{\sum_{i=1}^{d}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{d}(x_i-\bar{x})^2\sum_{i=1}^{d}(y_i-\bar{y})^2}}$

Fuzzy kNN (Schiaffino et al., 2016) is a variation of kNN, which involves fuzzy logic, allowing a sample to belong to multiple classes.

2.4. Support vector machine (SVM)

SVM are supervised machine learning models that chooses a hyperplane with the greatest possible margin between the hyperplane and any point in the training set for accurate classification (Russel and Norvig, 2010). This means that the maximum distance between the training vectors **x** and the hyperplane **w** should be maximized:

$$max_{\boldsymbol{w},b}\min||\boldsymbol{x} - \boldsymbol{x}_i|| : \boldsymbol{w}\cdot\boldsymbol{x} + b = 0, \quad i = 1,\cdots,m$$

This method can be used for both classification and regression. The disadvantages of SVM include its substantial computational complexity and its difficulty to handle noisy data.

2.5. Hidden Markov Model (HMM)

The HMM, commonly used in speech and language processing, is a probabilistic model that relates a sequence of observations to a sequence of hidden classes or hidden states that explain the observations (Jurafsky and Martin, 2017). The major limitation of HMM is that the probability of remaining in a state is represented as a geometric distribution, which may not be a valid assumption depending on the system modeled (Taghva, 2011). In a Hidden state Markov model (HsMM; a.k.a. explicit duration HMM or variable-duration HMM the state durations are represented (Taghva, 2011). HsMM has greater accuracy and specificity but is associated with a significant increase in computational time (Taghva, 2011).

However, the choice of type of machine learning model to apply is dependent on the data type available, resulting in differences in accuracy and impacts on performance trade-offs (Cabitza and Banfi, 2017).

3. Methods

A literature search was performed via databases such as Pubmed and Research gate utilizing the following key words "Parkinson's disease; deep brain stimulation, microelectrode recording, target localization, artificial intelligence, machine learning."

Ten studies relevant to the automation of deep brain stimulation surgery in Parkinson's disease were identified from 2006. They have been summarized in Table 1, arranged in chronological order, and will be discussed in accordance to their features and machine learning algorithms. We found one study which investigated specific features in the discrimination of the GPi nucleus, (Falkenberg et al., 2006a,b) however it was not included for evaluation as machine-learning techniques were not utilized.

Data obtained from each study include year of publication, number of patients, trajectories and MER recordings, length of recording, feature selection, machine learning methods and types of structures identified. The computational MER features utilized by the aforementioned studies for automatic detection and visualization of the structures have been summarized in Table 2. The performance results in terms of accuracy are presented in Table 3. All the studies, with the exception of Cagnan et al. (2011), utilized supervised machine learning models with expert defined features, and two studies (Cardona et al., 2012; Ciecierski et al., 2014) reported online real-time implementation of their algorithms.

4. Technical considerations

4.1. Signal pre-processing

The length of data chosen for analysis is critical in order to reduce the impact of noise (Falkenberg et al., 2006a,b). Most authors recommended waiting 2 s before recording (Moran et al., 2006; Zaidel et al., 2009) to reduce influence of electrode movement, except for Valsky et al. (2017) who commenced after 0.5 seconds. The subsequent recorded length of stationary electrode MER

Table 1
Summary of papers in chronological order.

First author	Moran	Zaidel	Wong	Cagnan	Chaovalitwongse	Cardona	Ciecierski	Rajpurohit	Schiaffino	Valsky
Year	2006	2009	2009	2011	2011	2012	2014	2015	2016	2016
Patients, n	27	21	27	48	16	nr	nr	26	8	81
Data analyzed										
Recordings, n	nr	nr		6064	nr	190	16,000	nr	1760	4526
Trajectories, n	36	56	43	258	nr	nr	nr	56	nr	131
Length, s	5	>5	4	nr	4	nr	nr	4	1	>4.5
Spike independent features										
Estimated distance to target (EDT)	✓	✓								✓
Basal amplitude				✓			✓		✓	
Signal kurtosis									✓	
Curve length (CL)			✓		✓			✓	✓	
Threshold			✓		✓				✓	
Peaks			✓		✓			✓	✓	
Average nonlinear energy (ANE)			✓		✓				✓	
Zero crossings			✓		✓			✓	✓	
Teager energy								✓		
Normalized root mean square (NRMS)	✓	✓	✓		✓			✓		✓
Noise mode								✓		
Power spectral density (PSD)		✓		✓		✓	✓			✓[a]
Spike dependent features										
Modified burst index (MBI)			✓		✓		✓	✓	✓	
Pause index (PI) and Pause ratio (PR)			✓		✓			✓	✓	
Compound firing rate (CFS)			✓	✓	✓		✓	✓	✓	
Mean spike amplitude differential			✓		✓			✓	✓	
Standard deviation of interspike intervals (ISI rms)			✓		✓	✓		✓	✓	
Mean spike trigger frequency									✓	
Spike fraction			✓		✓			✓		
Machine learning models										
Bayesian	✓				✓[b]	✓[c]	✓	✓[d]		
Decision tree					✓[e]		✓[f]			
Support vector machine								✓		✓
K-nearest neighbor			✓[g]		✓[h]			✓[i]	✓[g]	
Logistic regression								✓		
Hidden Markov model		✓[j]						✓		✓
Supervised learning	✓	✓	✓		✓	✓	✓	✓	✓	✓
Online						✓	✓			
Sites										
Thalamus								nr	✓	
Zona incerta					✓			nr		
Subthalamic nucleus entry	✓	✓	✓	✓	✓	✓	✓	nr	✓	
Subthalamic nucleus exit	✓	✓	✓	✓	✓			nr		✓
Substantia nigria					✓			nr	✓	

nr: not reported.
[a] Power ratio.
[b] Gaussian and kernel.
[c] Linear discriminant and quadratic classifier.
[d] Gaussian Naïve Bayes.
[e] Gini, max deviance.
[f] Random forest; rough set.
[g] Fuzzy.
[h] Euclidean, DTW.
[i] k = 5.
[j] Viterbi algorithm.

is suggested to be at least 4 s (Chaovalitwongse et al., 2011; Rajpurohit et al., 2015; Wong et al., 2009; Zaidel et al., 2009) for optimal tradeoff between feature stability and real-time processing.

Signals acquired during MER recordings arise from several sources as shown in Fig. 3. Neural activity in the form of neuronal spikes and background neural activity (He, 2009; Schiaffino et al., 2016) maybe contaminated by artifacts (noise) due to internal recording equipment (Lima, 2015; Snellings et al., 2009) or external environmental factors. (Lima, 2015) The presence of noise restricts the ability to obtain meaningful information from the signal (Cabitza and Banfi, 2017; Roy et al., 1999). Therefore preprocessing of signals is essential in order to achieve better performance for analysis. Noise artifacts may be detected by a combination of amplitude (Cagnan et al., 2011; Moran et al., 2006) and frequency criterion (Cagnan et al., 2011) relative to the estimated background neural activity. Other methods include that of wavelet transforming de-noising algorithms, first introduced by Donoho and Johnstone in 1995, (Novak et al., 2011; Snellings et al., 2009) and by rejecting MER segments with abnormal RMS values (Moran et al., 2006). Alternatively, neuronal oscillations for analysis may be identified via specific spike templates utilizing specific spike width and duration (Cagnan et al., 2011; Danish et al., 2008).

4.2. Feature normalization

Variable recording characteristics such as different recording lengths and trajectories, different amplifications, machine settings, different electrodes and inter-patient variation of neurophysiology result in a multidimensional set of properties (Ciecierski et al.,

Table 2
Microelectrode recording features.

Features	Description
Spike independent	
Estimated distance to target (EDT)	Estimated distance to target
Basal amplitude	Noise level estimation using envelope of MER signal using the Hilbert transform
Signal kurtosis	Signal kurtosis
Curve length (CL)	Sum of consecutive distances between points in the data per unit time
Threshold	Threshold in the signal amplitude
Peaks	Number of positive peaks in the data window per unit time
Average nonlinear energy (ANE)	Average energy difference between each signal sample and its two neighboring samples
Average absolute difference (AAD)	Average signal deviation to the mean signal amplitude
Zero crossings	The number of zero crossings in the data window per unit time.
Normalized root mean square (NRMS)	Normalized root mean square value of amplitude
Noise mode	Rajpurohit: 3 times the standard deviation of the amplitudes in the data window
Power spectral density (PSD)	Frequency response of a random/ periodic signal
Spike dependent	
Modified burst index (MBI)	Ratio of the number of interspike intervals less than10 ms to the number more than 10 ms
Pause index (PI) and Pause ratio (PR)	PI is the ratio of the count of ISIs that are greater than 50 ms to the number less than 50 ms; PR is the cumulative ISI time that greater than 50 ms divided by the cumulative ISI time less than 50 ms. Tonic and random firing have longer ISI so higher value of PI and PR were observed
Compound firing rate (CFS)	Number of spikes in one second
Mean spike amplitude differential	80% trimmed mean of the difference between consecutive spikes
Standard deviation of ISI (ISI rms)	Tonic firing – zero ISI RMS; Random firing – highest ISI RMS
Interspike intervals (ISI)	Three firing patterns: tonic, random and bursting. Tonic means the neuron fires periodically. Random means ISI follows a Poisson process
Spike fraction	Fraction percentage of spikes accepted as genuine spikes
Tonic index (TI)	Ratio of the summation of all of these shorter ISIs to the summation for all ISIs
Poisson surprise (PS)	Evaluating the probability of the occurrence of irregular events

2014). A comparable homogenous data set, especially for background neural attributes (Ciecierski et al., 2014; Taghva, 2011) is usually a pre-requisite for subsequent stages of analysis such as feature extraction, selection and classification (Chaovalitwongse et al., 2011). According to Rajpurohit (Rajpurohit et al., 2015), patient-specific normalization can improve classification performance by an average of 31.4% (Lima, 2015).

The caveat is that patient-specific normalization is typically determined only after the data has been collected and poorer accuracy may be observed if normalization parameters have not yet stabilized. Hence, this may present a problem in real-time implementations.

In contrast, Cardona et al. (2012) recommended against normalization of the signals, as accuracy of their system's algorithm decreased from 94% to 85% after normalization. The authors postulated that normalization affects the amplitudes of the recordings, resulting in a loss of relevant information from either the background neural activity or the high-frequency components.

4.3. Feature extraction and selection

Feature extraction is defined as the process of transforming a signal to retrieve specific properties (Lima, 2015). Combining different features for STN detection may increase the algorithm's accuracy (Cagnan et al., 2011; Wong et al., 2009), however, irrelevant or redundant features need to be removed to improve classification accuracy and computational efficiency as well as lower the number of required training instances for classifier training (Rajpurohit et al., 2015; Wong et al., 2012).

Rajpurohit et al. (2015) observed that out of 13 features, the top 1% of performing classifiers contained approximately 8.25 features on average. Chaovalitwongse et al. (2011) similarly concluded that not all 13 features tested in their algorithm were useful in the classification. This suggests that feature addition only improves performance up to a limit, likely due to feature redundancy or features that negatively contributed to classification.

5. Results

5.1. Spike-dependent features

Spike-dependent parameters are susceptible to errors and computational difficulties (Kim and McNames, 2007), especially when

Table 3
Accuracy of machine learning algorithms in identifying borders of STN and its surrounding structures.

Target	Study	Sensitivity	Specificity
Thalamus (TH)	Schiaffino 2016	0.86	0.90
Zona incerta (ZI)	Chaovalitwongse 2011	0.91	0.91
Subthalamic nucleus (STN)	Ciecierski 2014	0.93	0.98
	Chaovalitwongse 2011	0.91	0.89
	Cardona 2012	0.85	
	Schiaffino 2016	0.72	0.82
		Mean entry depth into STN	
	Wong 2009	−0.063 ± 0.44 mm	
	Zaidel 2009	−0.09 ± 0.35 mm	
	Moran 2006	−0.18 ± 0.84 mm	
	Cagnan 2011	88% agreement	
		Mean exit depth from STN	
	Valsky 2016	0.04 ± 0.18 mm	
	Wong 2009	−0.04 ± 0.31 mm	
	Zaidel 2009	−0.20 ± 0.33 mm	
	Moran 2006	0.50 ± 0.59 mm	
	Cagnan 2011	88% agreement	
Substantia nigra (SN)	Chaovalitwongse 2011	0.86	0.87
	Schiaffino 2016	0.78	0.75

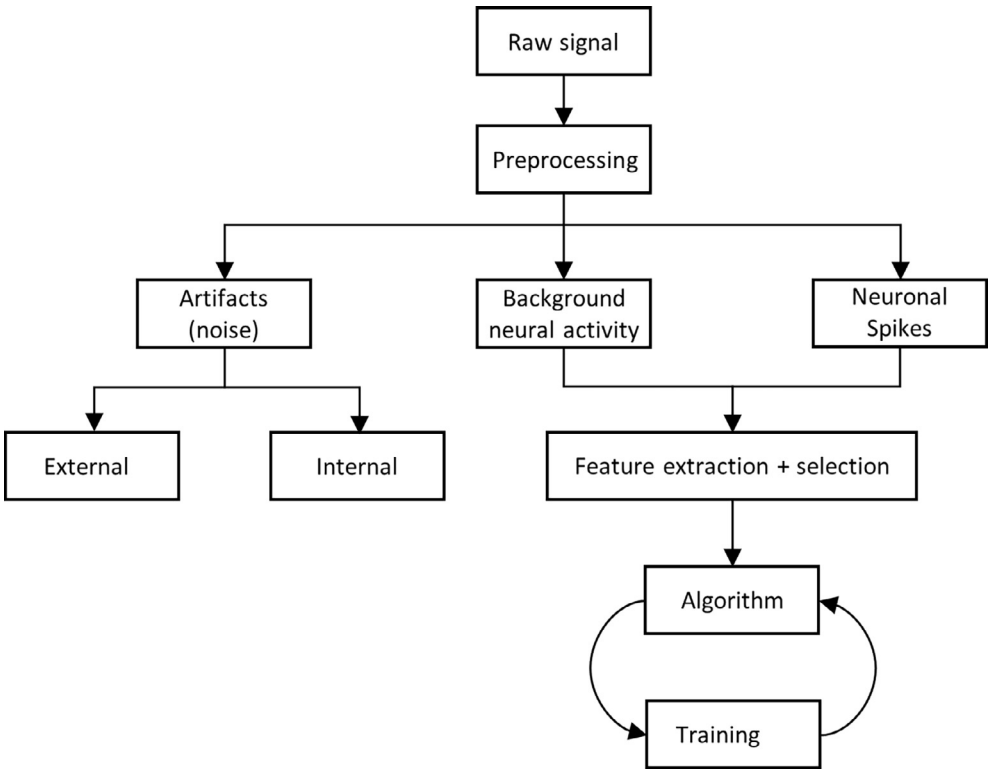

Fig. 3. Signal processing and machine learning. The raw MER signals are extracted and processed for analysis. As a first step, artifacts are detected and removed. The remaining data can be divided into background neural noise and spikes. Different features are extracted from this data, and features are selected for the machine-learning algorithm. This algorithm is trained on labeled MER data in order to improve its accuracy.

computed in real-time (Moran et al., 2006; Valsky et al., 2017). Stationary electrodes and longer recording times are required for both accurate spike detection and good signal-to-noise recording environments, thereby increasing the time and complexity of the procedure (Moran et al., 2006; Snellings et al., 2009; Wong et al., 2012). Therefore, there is a recent trend towards spike-independent features.

5.2. Spike-independent features

Both Moran et al. (2006) and Zaidel et al. (2009) reported a dramatic increase and decrease in normalized RMS (NRMS) with the entry and exit from STN respectively. Moran applied a Bayesian inference system based on two parameters: estimated distance to target (EDT) and NRMS of the electrical signal recorded by the electrode. NRMS is spike independent and is thus unaffected by variability in electrode properties, requiring only a short signal recording. The entry to STN could be predicted with a distinguishable rise in NRMS, however, the STN exit was less well predicted in some trajectories. Zaidel et al. (2009) subsequently improved upon this result by utilizing beta power spectral density (PSD) with a Hidden Markov model. They were able to distinguish the STN with better accuracy compared to the Bayesian probability as well as demarcate sensorimotor subterritory of STN with an error of -0.27 ± 0.58 mm. A recent publication by Valsky et al. (2017) improved upon the accuracy of STN exit detection with a novel idea of utilizing the power ratio between 100–150/5–25 Hz and a Hidden Markov Model. They achieved an impressive accuracy of 0.04 ± 0.18 mm (1 mm threshold) with a computing time of 99 ms.

An alternative method to NRMS was proposed by Novak et al. (2011), utilizing the multiunit activity (MUA) calculated in frequency domain. MUA is an aggregate of the spiking activity in the vicinity of the recording electrode (Logothetis, 2003). Novak et al. (2011) observed that MUA is elevated in the STN compared to its neighboring structures (Novak et al., 2007); it is also less sensitive to sudden changes in focal neuronal activity, since foreground spikes are removed to avoid spurious STN detection. Novak et al. (2011) achieved reasonably good correlation (STN dorsal/ventral borders 0.79, $p < 0.0001$/ 0.91, $p < 0.0001$) between MUA and intraoperative monitoring. However, MUA may still be contaminated by noise, interference and tremor - a common problem with all signal-processing methods. The authors were unable to analyze one subject's recordings due to low signal quality from severe electrical interference in the operating room (Novak et al., 2011).

5.3. Combination of spike-dependent, spike-independent, and background neural activity

Other authors (Cagnan et al., 2011; Chaovalitwongse et al., 2011; Ciecierski et al., 2014; Schiaffino et al., 2016) have applied a combination of spike-dependent features, background neural activity and/or spike-independent features for classification methods.

Background neural activity depends primarily on the quantity and activity of neurons in the vicinity of the electrode (Ciecierski et al., 2014) and is often detected using the noise mode envelope method (Cagnan et al., 2011; He, 2009).

One of the qualitative characteristics of STN electrophysiology is an increase in background activity of more than two-fold compared to zona incerta or internal capsule (Kano et al., 2008), likely due to an increased density of neurons within the STN (Benazzous et al., 2002; Snellings et al., 2009). Wavelet derived background amplitude levels, calculated after excluding foreground spikes, were found to be highest in regions encompassed by the clinically estimated boundaries of STN (Snellings et al., 2009).

This characteristic has been explored using four features: background curve length, RMS, amplitude and signal power (Ciecierski et al., 2014; Danish et al., 2008). With visual inspection of the background RMS feature profile, Danish et al. (2008) reported 90% sensitivity using a 0.5 mm precision criterion and 80% sensitivity with 0.5 mm precision criterion of identifying the STN entry and exit points respectively. The background curve length feature was found to have a much reduced sensitivity (60% for STN entry and 80% for STN exit, both with 0.5 mm precision criterion).

Cagnan et al. (2011) investigated the use of background neural activity, firing rate and PSD of low band index (3–12 Hz), beta band index (13–30 Hz) and gamma band index (31–100 Hz). All these features, except for low-band index, showed significantly higher values throughout the STN ($p < 0.001$). Background neural activity and firing rates were good indicators for the STN dorsal border, but less reliably so for the STN ventral border. Cagnan et al.'s classifiers achieved an 88% agreement with surgical annotations. However, out of 258 trajectories, the algorithm had 7 false positives due to low amplitude artifacts; as well as 8 false negatives, since the background neural activity remain below threshold and consecutive sites did not exhibit high firing rates.

Ciecierski et al. (2014) employed a combination of spike dependent features (spike frequency and modified burst ratio) and background neural attributes (absolute amplitude value, RMS, low frequency power in 0–500 Hz and high frequency in 500–3000 Hz). They achieved excellent accuracy of STN detection of 97.6% and 96.7% using Weka's Random Forest and Rough Set Exploration system respectively. They found that a combination of features had a significantly greater accuracy, sensitivity and specificity compared to using spike features alone, possibly due to false positive or negatives that may occur with spike discrimination. The reported computing time of two minutes for their algorithm makes it a possible option for intraoperative usage.

Similarly, Chaovalitwongse et al. (2011) was able to detect STN and its adjacent structures based on 7 spike detection-dependent and 6 spike detection-independent features, with almost 90% accuracy. Decision trees yielded the best performance with an accuracy of 87.6% and 89.6% when using Gini and Max Deviance indices respectively, compared to Bayesian (84.8%) and KNN (82.3%) models. This emphasizes that feature selection is important in subcortical classification, as both Bayesian and KNN models do not exploit any feature-selection methods within their framework. However, Schiaffino et al. (2016) demonstrated that their fuzzy KNN detection algorithm performed significantly better than the KNN algorithm ($p < 0.01$).

6. Conclusion

MER remains a common and widely used tool in DBS surgery, although there have been controversies regarding its use, (Falkenberg et al., 2006a,b; Hariz and Fodstad, 1999) especially with the increased interest in image guided DBS surgery (Cui et al., 2016; McEvoy et al., 2015; Ostrem et al., 2016). It is however evident that MER remains a useful adjunct for STN localization - an international survey conducted in 2010 revealed that 149 out of 185 DBS surgeons polled from 143 DBS centers still used microelectrode recording to obtain physiological confirmation of the target structure (Abosch et al., 2013). The key limitation of this review is the heterogeneity of the included studies. The studies had various MER acquisition formats and subsequently employed differing pre-processing techniques. The performances of the different classifiers were inconsistent in presentation and were reported either as error rates or sensitivity, specificity, and accuracy of their algorithms.

Nevertheless, the state-of-the-art features and methods for MER automation demonstrate delineation of the STN borders and its adjacent structures with good accuracy and relatively small error rates. Other possible areas for improvements may include the use of local field potentials (Bour et al., 2015; Chen et al., 2006; Telkes et al., 2014; Telkes et al., 2016), burst firing patterns to locate subterritories (Pozzi et al., 2016) and the use of hidden state Markov models (Taghva, 2011) to improve accuracy and computing time for automation. Promising semi-automated techniques such as the HaGuide Tool, have received FDA approval and are currently being validated in multiple groups (Thompson et al., 2018). Although both GPi and STN are potential sites for DBS in PD, there are no available studies regarding the automation of localization of the GPi nucleus. This highlights an area for further research to improve DBS surgical targeting and outcomes. With computational advancement, including the use of deep neural networks and the advent of deep feature synthesis algorithms (Kanter and Veeramachaneni, 2015), this will bring us closer towards developing real time, mathematically rigorous, fully automated techniques to analyze MER data (Shamir et al., 2015). These methods need to be accessible, cost-effective and time efficient, (Chibirova, 2006) hence thereby increasing the widespread availability of DBS surgery.

Funding

This research is supported by the Singapore Ministry of Health's National Medical Research Council (NMRC/CNIG/1173/2017).

Declarations of interest

None of the authors have potential conflicts of interest to be disclosed.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors. For this type of study formal consent is not required.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinph.2018.09.018.

References

Abosch A, Timmermann L, Bartley S, Rietkerk HG, Whiting D, Connolly PJ, et al. An international survey of deep brain stimulation procedural steps. Stereotact Funct Neurosurg 2013;91:1–11. https://doi.org/10.1159/000343207.

Benazzous A, Breit S, Koudsie A, Pollak P, Krack P, Benabid AL. Intraoperative microrecordings of the subthalamic nucleus in Parkinson's disease. Mov Disord 2002;17(3):145–9.

Bour LJ, Lourens MA, Verhagen R, de Bie RM, van den Munckhof P, Schuurman P, et al. Directional recording of subthalamic spectral power densities in Parkinson's disease and the effect of steering deep brain stimulation. Brain Stimul 2015;8(4):730–41. https://doi.org/10.1016/j.brs.2015.02.002.

Brocker DT, Swan BD, So RQ, Turner DA, Gross RE, Grill WM. Optimized temporal pattern of brain stimulation designed by computational evolution. Sci Transl Med 2017;9(371):1–11. https://doi.org/10.1126/scitranslmed.aah3532.

Cabitza F, Banfi G. Machine learning in laboratory medicine: waiting for the flood? Clin Chem Lab Med 2017;56(4):516–24.

Cagnan H, Dolan K, He X, Contarino MF, Schuurman R, van den Munckhof P, et al. Automatic subthalamic nucleus detection from microelectrode recordings based on noise level and neuronal activity. J Neural Eng 2011;8(4). https://doi.org/10.1088/1741-2560/8/4/046006 046006.

Cardona HDV, Padilla JB, Arango R, Carmona H, Alvarez MA, Estelles E, et al. NEUROZONE: On-line recognition of brain structures in stereotactic surgery – application to Parkinson's disease. Annual international conference of the IEEE engineering in medicine and biology society. IEEE Engineering in Medicine and Biology Society; 2012.

Chaovalitwongse WA, Jeong YS, Jeong MK, Danish SF, Wong S. Pattern recognition for identifying subcortical targets during deep brain stimulation surgery. Brain Inform 2011;26:2–11.

Chen CC, Pogosyan A, Zrinzo LU, Tisch S, Limousin P, Ashkan K, et al. Intra-operative recordings of local field potentials can help localize the subthalamic nucleus in Parkinson's disease surgery. Exp Neurol 2006;198(1):214–21. https://doi.org/10.1016/j.expneurol.2005.11.019.

Chibirova, O. Functional interactions in basal ganglia studied by multiple single unit recordings sorted by an unsupervised template matching algorithm. Neurons Cognit [q-bio.NC]. Université Joseph-Fourier - Grenoble I, 2006. English.; 2006 <tel-00115720>.

Ciecierski K, Mandat T, Rola R, Raś ZW, Przybyszewski AW. Computer aided subthalamic nucleus (STN) localization during deep brain stimulation (DBS) surgery in Parkinson's patients. Ann Acad Med Siles 2014;68(5):275–83.

Cui Z, Pan L, Song H, Xu X, Xu B, Yu X, et al. Intraoperative MRI for optimizing electrode placement for deep brain stimulation of the subthalamic nucleus in Parkinson disease. J Neurosurg 2016;124:62–9.

D'Haese PF, Pallavaram S, Li R, Remple MS, Kao C, Neimat JS, et al. CranialVault and its CRAVE tools: A clinical computer assistance system for Deep Brain Stimulation (DBS) therapy. Med Image Anal 2012;16(3):744–53. https://doi.org/10.1158/1078-0432.CCR-13-3047.

Danish SF, Baltuch GH, Jaggi JL, Wong S. Determination of subthalamic nucleus location by quantitative analysis of despiked background neural activity from microelectrode recordings obtained during deep brain stimulation surgery. J Clin Neurophysiol 2008;25(2):98–103. https://doi.org/10.1097/WNP.0b013e31816b38dd.

Falkenberg JH, McNames J, Burchiel KJ. Automatic microelectrode recording analysis and visualization of the globus pallidus interna and stereotactic trajectory. Stereotact Funct Neurosurg 2006a;84(1):28–34. https://doi.org/10.1159/000093720.

Falkenberg JH, McNames J, Favre J, Burchiel KJ. Automatic analysis and visualization of microelectrode recording trajectories to the subthalamic nucleus: Preliminary results. Stereotact Funct Neurosurg 2006b;84(1):34–44. https://doi.org/10.1159/000093721.

Friedman J. Greedy function approximation: a gradient boosting machine. IMS 1999 Reitz Lecture, 1999.

Gulberti A, Moll C, Hamel W, Buhmann C, Koeppen JA, Boelmans K, et al. Predictive timing functions of cortical beta oscillations are impaired in Parkinson's disease and influenced by L-DOPA and deep brain stimulation of the subthalamic nucleus. Neuroimage Clin 2015;9:436–49. https://doi.org/10.1016/j.nicl.2015.09.013.

Guo T, Parrent AG, Peters TM. Surgical targeting accuracy analysis of six methods for subthalamic nucleus deep brain stimulation. Comput Aided Surg 2007;12(6):325–34. https://doi.org/10.1080/10929080701730987.

Hariz M, Fodstad H. Do microelectrode techniques increase accuracy or decrease risks in pallidotomy and deep brain stimulation? A critical review of the literature. Stereotact Funct Neurosurg 1999;72:157–69.

He X. Neural signal processing of microelectrode recordings for deep brain stimulation Thesis. Sweden: Chalmers University of Technology Göteborg; .

Holt AB, Netoff TI. Computational modeling to advance deep brain stimulation for the treatment of Parkinson's disease. Drug Discov Today Dis Models 2016;19:31–6. https://doi.org/10.1016/j.ddmod.2017.02.006.

Jurafsky D, Martin J. Hidden markov models. Speech and Language processing. Stanford University; 2017.

Kano T, Katayama Y, Kobayashi K, Kasai M, Oshima H, Fukaya C, et al. Multiple-cell spike density and neural noise level analysis by semimicroelectrode recording for identification of the subthalamic nucleus during surgery for Parkinson's disease. Neuromodulation 2008;11(1):1–7.

Kanter J, Veeramachaneni K. Deep feature synthesis: Towards automating data science endeavors analytics (DSAA), 2015.

Karamintziou SD, Custódio AL, Piallat B, Polosan M, Chabardès S, Stathis PG, et al. Algorithmic design of a noise-resistant and efficient closed-loop deep brain stimulation system: A computational approach. PLoS One 2017;12(2):1–26. https://doi.org/10.1371/journal.pone.0171458.

Karamintziou SD, Deligiannis NG, Piallat B, Polosan M, Chabardès S, David O, et al. Dominant efficiency of nonregular patterns of subthalamic nucleus deep brain stimulation for Parkinson's disease and obsessive-compulsive disorder in a data-driven computational model. J Neural Eng 2016;13(1). https://doi.org/10.1088/1741-2560/13/1/016013 016013.

Kim S, McNames J. Automatic spike detection based on adaptive template matching for extracellular neural recordings. J Neurosci Methods 2007;165(2):165–74. https://doi.org/10.1016/j.jneumeth.2007.05.033.

Kühn AA, Volkmann J. Innovations in deep brain stimulation methodology. Mov Disord 2017;32(1):11–9. https://doi.org/10.1002/mds.26703.

Lima CFV. Analysis and classification of microelectrode recordings in deep brain stimulation surgery; 2015.

Logothetis NK. The underpinnings of the BOLD functional magnetic resonance imaging signal. J Neurosci 2003;23(10):3963–71.

Lourens MA, Meijer HG, Contarino MF, van den Munckhof P, Schuurman PR, van Gils SA. Functional neuronal activity and connectivity within the subthalamic nucleus in Parkinson's disease. Clin Neurophysiol 2013;124(5):967–81. https://doi.org/10.1016/j.clinph.2012.10.018.

Mallet L, Schupbach M, N'Diaye K, Remy P, Bardinet E, Czernecki V, et al. Stimulation of subterritories of the subthalamic nucleus reveals its role in the integration of the emotional and motor aspects of behavior. Proc Natl Acad Sci U S A 2007;104(25):10661–6.

McEvoy J, Ughratdar I, Schwarz S, Basu S. Electrophysiological validation of STN-SNr boundary depicted by susceptibility-weighted MRI. Acta Neurochir (Wien) 2015;157(12):2129–34. https://doi.org/10.1007/s00701-015-2615-1.

Moran A, Bar-gad I, Bergman H, Israel Z. Real-time refinement of subthalamic nucleus targeting using Bayesian decision-making on the root mean square measure. Mov Disord 2006;21(9):1425–31. https://doi.org/10.1002/mds.20995.

Novak P, Daniluk S, Ellias S, Nazzaro J. Detection of the subthalamic nucleus in microelectrographic recordings in Parkinson disease using the high-frequency (>500 Hz) neuronal background. J Neurosurg 2007;106:175–9.

Novak P, Przybyszewski AW, Barborica A, Ravin P, Margolin L, Pilitsis JG. Localization of the subthalamic nucleus in Parkinson disease using multiunit activity. J Neurol Sci 2011;310(1–2):44–9. https://doi.org/10.1016/j.jns.2011.07.027.

Okun MS, Tagliati M, Poufar M, Fernandez HH, Rodriguez RL, Alterman RL, et al. Management of referred deep brain stimulation failures. Arch Neurol 2005;62(8):1250–5.

Ostrem JL, Ziman N, Galifianakis NB, Starr PA, Luciano MS, Katz M, et al. Clinical outcomes using ClearPoint interventional MRI for deep brain stimulation lead placement in Parkinson's disease. J Neurosurg 2016;124:908–16. https://doi.org/10.3171/2015.4.JNS15173.

Pedregosa A. Scikit-learn: Machine Learning in Python. JMLR 2011;12:2825–30.

Pozzi NG, Arnulfo G, Canessa A, Steigerwald F, Nickl R, Homola GA, et al. Distinctive neuronal firing patterns in subterritories of the subthalamic nucleus. Clin Neurophysiol 2016;127(11):3387–93. https://doi.org/10.1016/j.clinph.2016.09.004.

Rajpurohit V, Danish SF, Hargreaves EL, Wong S. Optimizing computational feature sets for subthalamic nucleus localization in DBS surgery with feature selection. Clin Neurophysiol 2015;126(5):975–82. https://doi.org/10.1016/j.clinph.2014.05.039.

Rowland NC, Sammartino F, Lozano AM. Advances in surgery for movement disorders. Mov Disord 2017;32(1):5–10. https://doi.org/10.1002/mds.26636.

Roy M, Kumar VR, Kulkarni BD, Sanderson J, Rhodes M, Stappen M. Simple denoising algorithm using wavelet transform. AIChE J 1999;45(11):2461–6.

Russel SJ, Norvig P. Artificial intelligence a modern approach. https://doi.org/10.1017/S0269888900007724.

Schiaffino L, Rosado Munoz A, Guerrero Martinez J, Francés Villora J, Gutiérrez A, Martinez Torres I, et al. STN area detection using K-NN classifiers for MER recordings in Parkinson patients during neurostimulator implant surgery. J Phys Conf Ser 2016;705(1). https://doi.org/10.1088/1742-6596/705/1/012050.

Senders JT, Staples PC, Karhade AV, Zaki MM, Gormley WB, Broekman MLD, et al. Machine learning and neurosurgical outcome prediction: a systematic review. World Neurosurg 2017;109:476–86. https://doi.org/10.1016/j.wneu.2017.09.149.

Shamir RR, Dolber T, Noecker AM, Walter BL, McIntyre CC. Machine learning approach to optimizing combined stimulation and medication therapies for Parkinson's disease. Brain Stimul 2015;8(6):1025–32. https://doi.org/10.1126/science.aaa1823.Using.

Shamir RR, Zaidel A, Joskowicz L, Bergman H, Israel Z. Microelectrode recording duration and spatial density constraints for automatic targeting of the subthalamic nucleus. Stereotact Funct Neurosurg 2012;90(5):325–34. https://doi.org/10.1159/000338252.

Snellings A, Sagher O, Anderson DJ, Aldridge JW. Identification of the subthalamic nucleus in deep brain stimulation surgery with a novel wavelet-derived measure of neural background activity. J Neurosurg 2009;111(4):767–74. https://doi.org/10.1016/j.bbamem.2015.02.010.Cationic.

Taghva A. Hidden semi-Markov models in the computerized decoding of microelectrode recording data for deep brain stimulator placement. World Neurosurg 2011;75(5–6):758–763.e4. https://doi.org/10.1016/j.wneu.2010.11.008.

Telkes I, Ince N, Onaran I, Abosch A. Localization of subthalamic nucleus borders using macroelectrode local field potential recordings. In: Annual international conference of the IEEE engineering in medicine and biology society. IEEE Engineering in Medicine and Biology Society; 2014. p. 2621–4. https://doi.org/10.1109/EMBC.2014.6944160.

Telkes I, Jimenez-Shahed J, Viswanathan A, Abosch A, Ince NF. Prediction of STN-DBS electrode implantation track in Parkinson's disease by using local field potentials. Front Neurosci 2016;10(198):1–16. https://doi.org/10.3389/fnins.2016.00198.

Teplitzky BA, Zitella LM, Xiao YZ, Johnson MD. Model-based comparison of deep brain stimulation array functionality with varying number of radial electrodes and machine learning feature sets. Front Comput Neurosci 2016;10(58):1–15. https://doi.org/10.3389/fncom.2016.00058.

Thompson JA, Oukal S, Bergman H, Ojemann S, Hebb AO, Hanrahan S, et al. Semi-automated application for estimating subthalamic nucleus boundaries and optimal target selection for deep brain stimulation implantation surgery. J Neurosurg 2018;1–10. https://doi.org/10.3171/2017.12.JNS171964.

Valsky D, Marmor-Levin O, Deffains M, Eitan R, Blackwell KT, Bergman H, et al. Stop! border ahead: Automatic detection of subthalamic exit during deep brain stimulation surgery. Mov Disord 2017;32(1):70–9. https://doi.org/10.1002/mds.26806.

van Wijk BCM, Pogosyan A, Hariz MI, Akram H, Foltynie T, Limousin P, et al. Localization of beta and high-frequency oscillations within the subthalamic nucleus region. Neuroimage Clin 2017;16:175–83. https://doi.org/10.1016/j.nicl.2017.07.018.

Williams NR, Foote KD, Okun MS. Subthalamic nucleus versus globus pallidus internus deep brain stimulation: translating the rematch into clinical practice. Mov Disord Clin Pract 2014;1(1):24–35. https://doi.org/10.1002/mdc3.12004.

Wong S, Baltuch GH, Jaggi JL, Danish SF. Functional localization and visualization of the subthalamic nucleus from microelectrode recordings acquired during DBS surgery with unsupervised machine learning. J Neural Eng 2009;6(2). https://doi.org/10.1088/1741-2560/6/2/026006.

Wong S, Hargreaves EL, Baltuch GH, Jaggi JL, Danish SF. Depth-time interpolation of feature trends extracted from mobile microelectrode data with kernel functions. Stereotact Funct Neurosurg 2012;90(1):51–8. https://doi.org/10.1159/000334494.

Zaidel A, Spivak A, Shpigelman L, Bergman H, Israel Z. Delimiting subterritories of the human subthalamic nucleus by means of microelectrode recordings and a hidden Markov model. Mov Disord 2009;24(12):1785–93. https://doi.org/10.1002/mds.22674.