# DTI Helps to Predict Parkinson's Patient's Symptoms Using Data Mining Techniques

Artur Chudzik[1(✉)], Artur Szymański[1], Jerzy Paweł Nowacki[1], and Andrzej W. Przybyszewski[1,2]

[1] Polish-Japanese Academy of Information Technology, Koszykowa 86 St., 02-008 Warsaw, Poland
{artur.chudzik, artur.szymanski, nowacki, przy}@pjwstk.edu.pl

[2] Department of Neurology, University of Massachusetts Medical School, 65 Lake Avenue, Worcester, MA 01655, USA
andrzej.przybyszewski@umassmed.edu

**Abstract.** Deep Brain Stimulation (DBS) is commonly used to treat, inter alia, movement disorder symptoms in patients with Parkinson's disease, dystonia or essential tremor. The procedure stimulates a targeted region of the brain through implanted leads that are powered by a device called an implantable pulse generator (IPG). The mentioned targeted region is mainly chosen to be subthalamic nucleus (STN) during most of the operations. STN is a nucleus in the midbrain with a size of 3 mm × 5 mm × 9 mm that consist of parts with different physiological functions. The purpose of the study was to predict Parkinson's patient's symptoms defined by Unified Parkinson's Disease Rating Scale (UPDRS) that may occur after the DBS treatment. Parameters had been obtained from 3DSlicer (Harvard Medical School, Boston, MA), which allowed us to track connections between the stimulated part of STN and the cortex based on the DTI (diffusion tensor imaging).

**Keywords:** Subthalamic nucleus · UPDRS · RSES · MRI · DTI · DBS · Parkinson's disease · Data mining

## 1 Introduction

Neurodegenerative diseases, in which we could distinguish Parkinson Disease (PD), have their background in neurodegeneration which could be described as progressive loss of structure or function of neurons, including the death of neurons. PD is primarily related to the substantia nigra degeneration which leads to dopamine insufficiency. Standard medication in PD is L-DOPA, which is a precursor of dopamine. However, disease progression affects in L-DOPA efficiency decay which may be revealed in on-off symptom fluctuation.

Thus, the neurologist has often to extend standard medication therapy to DBS (Deep Brain Stimulation) surgery [1]. DBS treatment depends on stimulation of the subthalamic nucleus (STN) which is dorsal to the substantia nigra and medial to the internal capsule. STN is also being known as a "hyper direct pathway" [2] of motor


N. T. Nguyen et al. (Eds.): ACIIDS 2019, LNAI 11432, pp. 615–623, 2019.
https://doi.org/10.1007/978-3-030-14802-7_53

control, contrasting with the direct and indirect pathways implemented elsewhere in the basal ganglia. However, the procedure of application the DBS electrode under the appropriate placement is challenging and may affect in different recovery time and treatment effectiveness.

The searching of localization of the subthalamic nucleus is done mainly by the registrations of neuronal activity via microelectrode recording (MER) [8]. MER is an intraoperative analysis of multi-unit activity (MUA). The commonly used criteria for electrophysiological localization of the STN are qualitative and mainly based on visual and acoustic observations of changes in spike frequencies and background activity. The characteristics of spike trains change during the whole path of brain structures and differ when the electrode passes through the thalamus, zona incerta, lenticular fasciculus, subthalamic nucleus, and the substantia nigra. Bursts in the background activity and sudden increases in the frequency of neuronal spiking are signs that electrode is near to STN. To obtain additional confirmation of the correct electrode placement, supplementary kinesthetic responses measurements aligned with microstimulation are being proceeded. There are two main strategies in searching for STN. First one depends on a single microelectrode which leads to the necessity of multiple passes for correct localization of the motor region. The second one uses 4–5 microelectrode insertions simultaneously. It has to be noticed that any stimulation or manipulation of the non-motor STN region is usually avoided since it can provoke psychiatric and cognitive dysfunctions [7].

To predict the neurological effects related to different electrode-contact stimulations, we have extracted specific parameters acquired from diffusion tensor imaging (DTI). We have demonstrated that with the data mining methods, supported with the rough set theory, it is possible to predict Parkinson's patient's symptoms, according to Unified Parkinson's Disease Rating Scale (UPDRS) [9].

## 2 Methods

In this research, the subject of study was data acquired from nine patients with advanced PD, which have had DBS electrodes implanted. The primary step was the analysis of the data acquired from the DTI by 3DSlicer software. Those parameters were: two technical values (fiducial region size which determines the tractography radius for selected electrode contact; stopping value - the value of ceasing for the generation of the given tract) and an amount of tract reaching the proximity of given ROI, distinguished between left and right side for every region.

The process of the tractography generation was described in previous works [3, 4]. The generation was carried separately for each contact, and it was on DTI data from the pre-OP DWI (pre-operational diffusion-weighted imaging). The DWI to DTI (diffusion tensor imaging) data was estimated by the use of least squares function approximation. Then, to generate relevant tracts, a proper ROI (region of interest) has been set for each patient, based on electrode contacts. Next, a module called Tractography Interactive Seeding has been applied in order to generate tracts. For every patient, two sets of data have been generated- with a minimal and large (over 30) number of tracts into primary and supplementary motor cortices (Fig. 1). The parameters that were used during the creation of individual tractography were fiducial region size and stopping value, mentioned previously.

The analysis included discovering of the correlation between given attributes aligned with the importance. This operation was performed with the usage of *pandas* library, which is a Python tool for data analysis and statistics [6]. Based on the obtained values, it was possible to select the attributes relevant in the data mining process, which was performed in RSES software.

We have used the RSES 2.1.1 (Rough System Exploration Program) with implementation of RS rules to process our data. An information system [5] can be considered as a pair:

$$S = (U, A)$$

Where:

$U = universe\ of\ objects$
$A = set\ of\ attributes$
$V = set\ of\ values$
$a(u) = unique\ element\ of\ V$
$a \in A$
$u \in U$
A decision table for $S$ is the 3-tuple:

$$S = (U, C, D)$$

Where:

$C = condition\ attributes$
$D = decision\ attributes$

Information table contains rows, where each denotes a particular rule that connects condition and decision attributes for a single measurement of a specific patient.

For results evaluation, we have used a technique called cross-validation, which is a suitable method for estimating the performance of a predictive model, selection of features or parameters adjustment. It is based on the approach of partitioning a data set $X$ into $n$ subsets $X_i$. Then, given algorithm is performed $n$ times, each time using a different training set $X - X_i$ and validating the results on $X_i$.

The classifier could be considered to be relevant because of attributes selection that had been done by an algorithm itself. For example, when the left hand tremor is taken as a prediction value, the RSES assumed that relevant attributes are related to, inter alia, hand tracts.

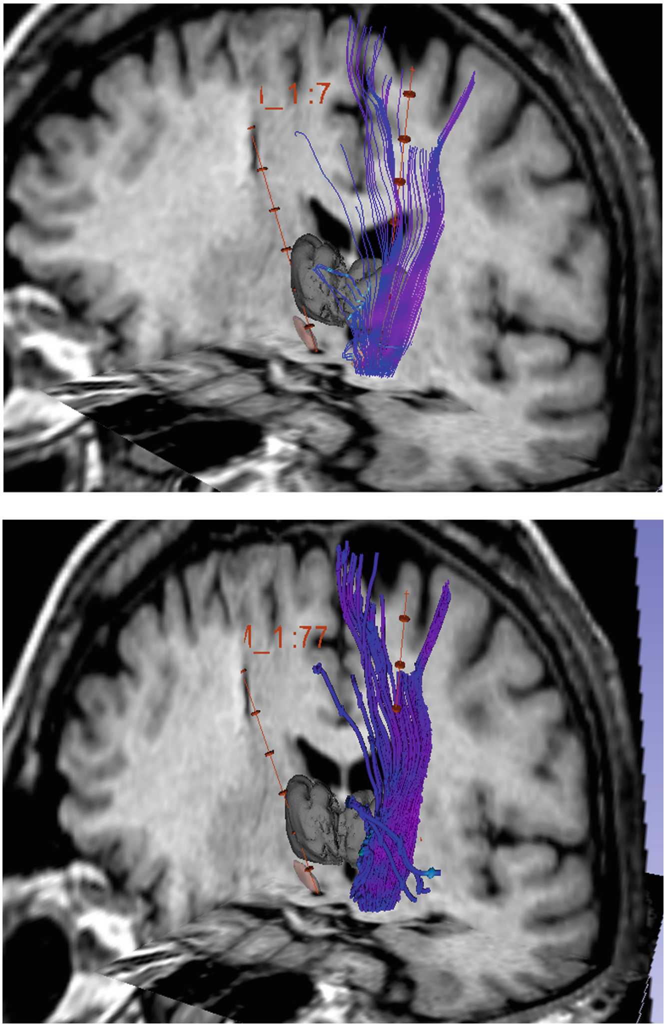

**Fig. 1.** Screenshots taken from Slicer 3D project of sample patient. We can observe trajectories of both implanted electrodes marked by orange lines, which are aligned with dead tissue visible on MRI slices as a result of surgery. Position of the electrodes is visualized with relevant neighboring structures: STN, Globus Pallidus and Thalamus. DTI tractography is generated from left STN showing connections going to SM and M1 areas of cortex. (Color figure online)

## 3 Results

The first step was to create a decision table that consists of data attributes obtained from 3DSlicer, based on diffusion tensor imaging, as described previously in the section on Methods. Then, RSES methods were applied, to get decision rules. To achieve that, rows and columns of Table 1 must have been exchanged so that parameters of different patients were in rows, and their results (attributes) were in columns.

**Table 1.** A fragment of the input dataset. UPDRS <code> - UPDRS value for specific motoric classification of patient' condition; Slicer L/R fiducial region size – Slicer tractography radius for selected electrode contact (in millimeters); Slicer L/R stopping value – Slicer parameter for ceasing generation of the given tract; Slicer L/R tracts lip/hand/foot – number of tract reaching proximity of lip/hand/foot ROI.

| Patient ID | 10 | 10 | 20 |
|---|---|---|---|
| UPDRS 21 L Hand action or postural tremor | 0 | 1 | 2 |
| Slicer R fiducial region size | 5 | 5 | 5 |
| Slicer R stopping value | 0.21 | 0.21 | 0.21 |
| Slicer R tracts hand | 2 | 3 | 3 |
| Slicer R tracts lip | 4 | 15 | 2 |
| Slicer R tracts foot | 15 | 2 | 2 |

In all experiments, we had used the Unified Parkinson Disease Rating Scale (UPDRS) which gave us information about the disease progression in various parts of the body in the context of disease-dependent factor (tremor).

The dataset has been spited into three, smaller subsets, where each was targeted to a different UPDRS marker ("left-hand tremor", "face, lips, chin tremor", "handwriting distortion"). On every subset, we had conducted an experiment based on defined attributes that led to the conclusion of possible UPDRS value after application of DBS treatment on the patient.

### 3.1 Prediction of the Left-Hand Tremor

Attributes that were relevant during the discretization process were strictly related to Slicer data acquired from the right side of the area, such as fiducial region size, stopping value, tracts for hand, lip and foot (Fig. 2).

The results of the prediction of the left-hand tremor based on given DTI parameters revealed the accuracy of 0.967 with the coverage of 0.425 (Table 2).

### 3.2 Prediction of the Facial Area Tremor

For this task, chosen attributes were related both to the left and right part of the area. From the left side: fiducial region size, stopping value and lip tracts. From the right side: stopping values, and tracts for lips and foot (Fig. 3).

Prediction result of the "face, chin, lips tremor" was with the accuracy as high as 0.824 with the coverage of 0.7 (Table 3).

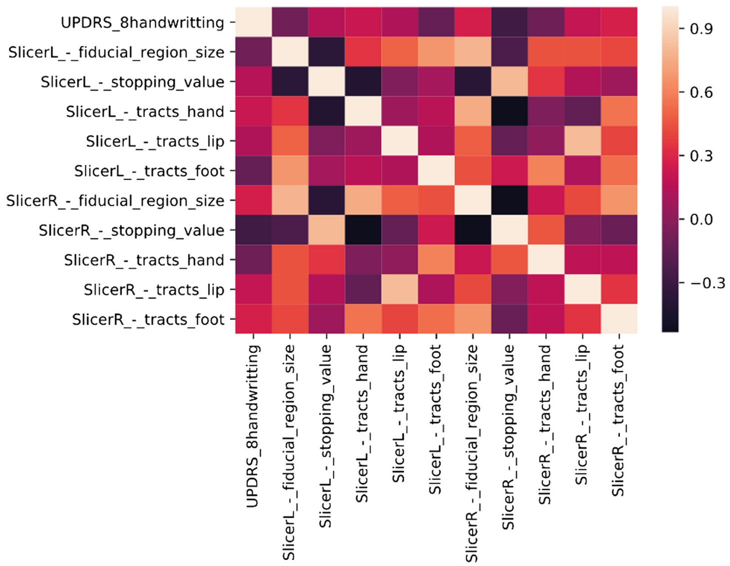


**Fig. 2.** Computed correlation between Slicer parameters and UPDRS examination: action or postural tremor of Hands – Left hand. The similarity between the two parameters is correlated with the color of the cell in the matrix (lighter shade represents higher similarity and contrariwise). (Color figure online)

**Table 2.** Confusion matrix of UPDRS #21: action or postural tremor of hands - left hand by rules obtained from 3DSlicer values based on DTI data. Number of tested subjects: 8. Accuracy (total): 0.967. Coverage (total): 0.425. TPR stands for "true positive rate".

| | Predicted | | | | | | |
|---|---|---|---|---|---|---|---|
| Actual | | **0** | **1** | **2** | **No. of obj.** | **Accuracy** | **Coverage** |
| | **0** | 2.6 | 0.0 | 0.0 | 5.2 | **0.8** | 0.48 |
| | **1** | 0.2 | 0.6 | 0.0 | 2.0 | **0.5** | 0.30 |
| | **2** | 0.0 | 0.0 | 0.0 | 0.8 | **0.0** | 0.00 |
| | **TPR** | **0.76** | **0.6** | **0.0** | | | |

### 3.3 Prediction of the Handwriting Disturbances

The last experiment with the detection of UPDRS value change of the "handwriting" test was with the accuracy of 0.878 by the coverage 0.667 (Fig. 4).

The relevant attributes were: DBS and BMT. Furthermore, mainly parameters from the left side Slicer were observed as relevant, such as fiducial region size, stopping value, hand and lip tracts. From the right side, only one technical value (stopping) was taken under consideration (Table 4).

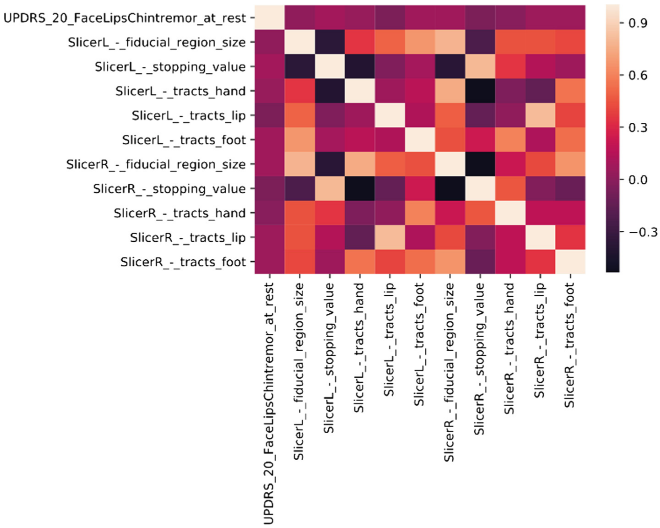


**Fig. 3.** Computed correlation between Slicer parameters and UPDRS examination: tremor at rest, face, lips, and chin. The similarity between the two parameters is correlated with the color of the cell in the matrix (lighter shade represents higher similarity and contrariwise). (Color figure online)

**Table 3.** Confusion matrix of UPDRS #20: tremor at rest, face, lips, chin by rules obtained from 3DSlicer values based on DTI data. Total number of tested subjects: 10. Accuracy (total): 0.824. Coverage (total): 0.7.

| | Predicted | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Actual | | **0** | **1** | **2** | **3** | **No. of obj.** | **Accuracy** | **Coverage** |
| | **0** | 4.75 | 0.50 | 0.00 | 0.25 | 6.5 | **0.867** | 0.866 |
| | **1** | 0.25 | 0.75 | 0.00 | 0.00 | 1.5 | **0.750** | 0.833 |
| | **2** | 0.00 | 0.00 | 0.00 | 0.00 | 0.5 | **0.000** | 0.000 |
| | **3** | 0.25 | 0.00 | 0.00 | 0.25 | 1.5 | **0.125** | 0.250 |
| | **TPR** | **0.93** | **0.50** | **0.00** | **0.25** | | | |

## 4 Discussion

Deep brain stimulation is currently widely applied as a surgical choice of treatment for patients with advanced PD. The benefits of STN stimulation are due to combined mechanisms and involve several adjacent structures. To improve the success of the procedure, more selectivity is needed and both topographical level and stimulation parameters must be enhanced [1].

This article represents the continuation of previous findings presented in [7] that are useful to the surgeon as a tool for confirmation that the subthalamic nucleus is near to the microelectrode path. Furthermore, it extends them even more with data mining techniques to predict the neurological effects related to different electrode-contact stimulations.

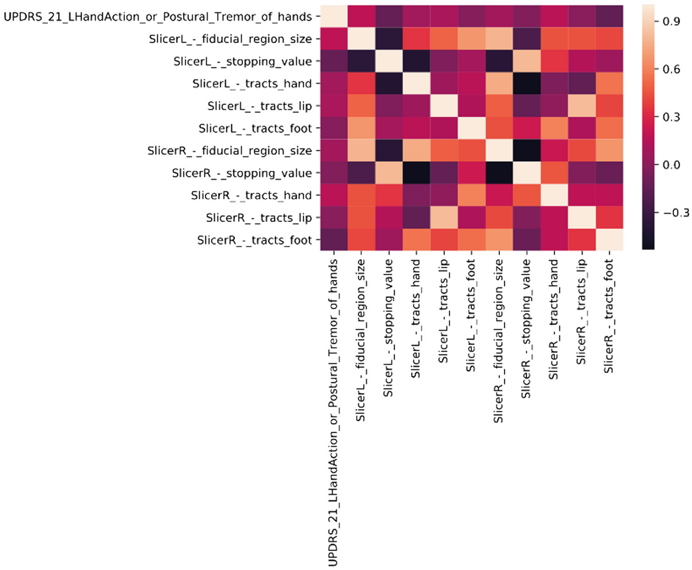


**Fig. 4.** Computed correlation between Slicer parameters and UPDRS examination: handwriting. The similarity between the two parameters is correlated with the color of the cell in the matrix (lighter shade represents higher similarity and contrariwise). (Color figure online)

**Table 4.** Confusion matrix of UPDRS #8: handwriting distortion by rules obtained from 3DSlicer values based on DTI data. Total number of tested subjects: 6. Accuracy (total): 0.878. Coverage (total): 0.667.

| | | Predicted | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **0** | **1** | **2** | **3** | **4** | **No. of obj.** | **Accuracy** | **Coverage** |
| Actual | **0** | 3.33 | 0.17 | 0.00 | 0.00 | 0.00 | 4.00 | **0.944** | 0.878 |
| | **1** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.83 | **0.000** | 0.000 |
| | **2** | 0.00 | 0.00 | 0.17 | 0.00 | 0.00 | 0.50 | **0.167** | 0.083 |
| | **3** | 0.00 | 0.00 | 0.33 | 0.00 | 0.00 | 0.33 | **0.000** | 0.167 |
| | **4** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.33 | **0.000** | 0.000 |
| | **TPR** | **1.00** | **0.00** | **0.17** | **0.00** | 0.00 | | | |

## 5 Conclusions

Our recent research described above was meant to determine if data mining can predict possible Parkinson's patient's symptoms based only on the DTI data of patients who go through the DBS surgery. We have applied the rough set theory on the data obtained from DTI after the operation to conclude whether it is possible to create a system that is unbiased of human opinion.

The results have shown that it is possible to introduce a new, autonomous, doctor independent and a highly accurate method of disease course prediction.

What is more, this approach enables a new way for a deduction of an impact of a region-specific stimulation of STN and its effect on patients.

However, since this results have been based on a small data set, further work is required to perform more credible statistics and verification in the sake of elimination of overfitting problem.

## References

1. Benabid, A.L., et al.: Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. Lancet Neurol. **8**(1), 67–81 (2009)
2. Nambu, A., Tokuno, H., Takada, M.: Functional significance of the cortico–subthalamo–pallidal 'hyperdirect' pathway. Neurosci. Res. **43**(2), 111–117 (2002)
3. Szymański, A., Przybyszewski, A.W.: Rough set rules help to optimize parameters of deep brain stimulation in Parkinson's patients. In: Ślęzak, D., Tan, A.-H., Peters, J.F., Schwabe, L. (eds.) BIH 2014. LNCS (LNAI), vol. 8609, pp. 345–356. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-09891-3_32
4. Szymański, A., Kubis, A., Przybyszewski, A.W.: Data mining and neural network simulations can help to improve deep brain stimulation effects in Parkinson's disease. Comput. Sci. **16**(2), 199 (2015)
5. Pawlak, Z.: Rough set theory and its applications. J. Telecommun. Inf. Technol., 7–10 (2002)
6. McKinney, W.: Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference, vol. 445 (2010)
7. Przybyszewski, A.W., et al.: Multi-parametric analysis assists in STN localization in Parkinson's patients. J. Neurol. Sci. **366**, 37–43 (2016)
8. Benazzouz, A., et al.: Intraoperative microrecordings of the subthalamic nucleus in Parkinson's disease. Mov. Disord. Official J. Mov. Disord. Soc. **17**(S3), S145–S149 (2002)
9. Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease: The unified Parkinson's disease rating scale (UPDRS): status and recommendations. Mov. Disord. **18**(7), 738–750 (2003)