Building Classifiers for Parkinson's Disease Using New Eye Tribe Tracking Method

Artur Szymański[1(✉)], Stanisław Szlufik[2], Dariusz M. Koziorowski[2], and Andrzej W. Przybyszewski[1]

[1] Polish-Japanese Academy of Information Technology, Koszykowa 86, 02-008 Warsaw, Poland
{artur.szymanski,przy}@pja.edu.pl
[2] Faculty of Health Science, Department of Neurology, Medical University of Warsaw, Warsaw, Poland
stanislaw.szlufik@gmail.com, dkoziorowski@esculap.pl

Abstract. Parkinson Disease (PD) is the second major neurodegenerative disease, which causes severe complications for patients' daily life. PD remains unspecified in many aspects including best treatment, prediction of its progression and precise diagnosis. In our study we have built machine learning (ML) models, which address some of those issues by helping to improve symptom evaluation precision by using advanced biomarkers such as fast eye movements. We have built and compared model accuracy relaying on data from two systems for recording eye movements: one is saccadometer (Ober Consulting), and another is based on the Eye Tribe (ET1000). We have reached 85% accuracy in prediction of neurologic attributes based on ET and 82% accuracy with saccadometer with help of rough set theory. The purpose of this study was to compare ET with clinically approved eye movement measurements saccadometer of Ober. We have demonstrated in 8 PD patients that both systems gave comparable results based on neurological and eye movement measurements attributes.

Keywords: Data mining · Eye tracking · Parkinson Disease · Rough set theory

1 Introduction

As the one of the most common neurodegenerative diseases we still obtain a lot of imprecise diagnosis of Parkinson's Disease. Currently in treatment we rely heavily on experience of neurologist. Symptoms and disease progression can vary significantly between patients and it is unclear what is exactly optimal treatment. In our study we have used ML methods to build classifiers in order to assist in objective assessment of PD patients using reflexive saccades (RS) as biomarker.

Approaches of using ML in PD assessment have already been carried out, using variety of biomarkers to improve objective evaluation. For example, in work of Tsanas et al. [1] we find examples of using speech signal processing which can detect dysphonia in PD patients with accuracy reaching 99%. Other studies [2] shows examples of using machine learning on MRI data in order to classify PD clinically diagnosed

N.T. Nguyen et al. (Eds.): ACIIDS 2017, Part II, LNAI 10192, pp. 351–358, 2017.
DOI: 10.1007/978-3-319-54430-4_34

patients against control group. By analysing voxel data and processing it with classifier based on Support Vector Machines (SVM) it was possible to reach specificity and sensitivity above 90%.

Other recent studies [3] show that correlation of few biomarkers can give excellent results for prediction of early stages of PD. Also highlighting importance of such effort in treatment. Authors have shown that combination of non-motor features can provide high accuracy for predicting early PD. Using data from Parkinson's Progression Markers Initiative they were able to benchmark few classification approaches reaching 96% accuracy with SVM.

Our own efforts in building automated and doctor independent solution shows that ML approach could be extremely efficient in classification of PD and help neurologists in patient assessment. We have used different biomarkers to demonstrate their importance in PD diagnosis, including DTI imaging [4] and single-photon emission computed tomography [5]. In most recent works [6, 7] we have shown than RS data can be used for building intelligent classifiers which reach over 90% accuracy in predicting PD patient features which make them important biomarker.

In this work we present new approach of recording eye movement data using software developed by our team based on Eye Tribe (ET) framework. We have built models using data from two eye tracking systems: clinically approved saccadometer with a new in the clinic ET. We have demonstrated that a low cost ET framework can be effectively used in the clinic in order to improve prediction of PD symptoms.

2 Methods

In our study we analysed data of 8 PD patients, in 21 sessions. Every patient had from 1 to 4 sessions. For each patient standard neurological tests were recorded. Each session determined whether a particular patient has deep brain stimulation (DBS) of subthalamic nucleus or the best medical treatment (BMT) enabled. Patients tested: in session one (S1) were off DBS and off BMT; in session two (S2) were on DBS and off BMT; in session three (S3) were off DBS and on BMT; in session four (S4) were both on DBS and on BMT. Not all patients in our study have recordings for all four sessions examinations. As a qualification parameter we took quality of eye data captured for given patient, for given session. Eye data used in this study includes reflexive saccades (RS) which were recorded using two systems, one from Ober Consulting - saccadometer and another one developed by us on the bases of Eye Tribe (ET) tracker.

During procedure with Ober saccadometer, a patient sat in front of the wall with the device mounted on his/her head. After starting the procedure the patient saw a red dot in front of him/her. The dot moved randomly to the left or right, and after about a second came back to the central position. Patient's task was to follow fast moving spot, which is equivalent with performing RS. This experimental protocol was the same for both devices. By means of Ober saccadometer, data were recorded with sampling frequency of 1000 Hz.

The ET system has used infrared camera positioned in the front of the patient and under LCD monitor. Camera tracked positions of each eye separately. Before each examination patient was asked to perform calibration by performing short fixations on

7 or 9 spots displayed on the screen one by one in different locations. Following the calibration, the ET procedure were similar to those in Ober device with the difference that marker in ET set-up was displayed on the LCD screen. Data in ET were sampled with frequency of 30 Hz.

The process in ET solution is managed with Java application using ET API, providing build in functions for calibration, predefined and custom procedures for RS and pursuit eye movements, online preview of current procedure, simplified error correction and data preview module.

There were following differences in both systems: different stimulus displayed means, different sampling frequencies and different data presentation methods. From ET we are receiving signal data with help of provided framework while Ober produce aggregated static parameters like delay or latency of eye movement averaged for all saccades. ET method of displaying data is easier to access the raw signal data from which we can remove artefacts and manually process different data parameters. Signal data samples from ET are sent to our program every 1 s.

There is another important difference related to placing of these devices. Ober saccadometer is fixed on patient's head, while ET camera is positioned in front of patient under the LCD screen. These differences result in different sensitivities to artefacts related to patient's head movements.

Another, mentioned above difference is related to the light stimulus display. Ober has used a red dot with static spot location, which was jumping by 10-degree to the left or the right. In ET we use a light spot displayed on LCD screen. Movements of the spot on the screen were described in details above. Ober saccadometer has software that automatically calculate saccades attributes and is taking distance from the eye to the spot into account. In ET it was necessary to measure this distance for each subject that has strong influence max saccades speed calculations.

Data from ET solution displays simultaneously eye movement measurements and related movements of the light spot on the screen. We have analysed recorded data by rewriting algorithms from our previous studies [6] (with help of python with a standard frameworks like "numpy"). Attributes which we took into consideration included latencies, max speed and amplitude. All of them described in our previous studies [6].

Both systems had possibilities to measure movement of each eyes separately. However, in this study we did not take asymmetric eye responses into account, therefore we have averaged measurements for both eyes. Figure 1 represents an example of our saccades recording.

In addition to eye movements, we have collected standard neurological parameters for PD patients like age, sex, Unified Parkinson's Disease Rating, Hoehn and Yahr scale, Schwab England scale, PDQ (quality of life measurement) and others. Full index of used attributes is presented in the input data table in the results section.

We have processed our data with help of the machine learning and data mining software in order to build classifiers for predicting effects of different treatments (session number) and total UPDRS. We have used two data mining software: Rough Set Exploration System [8–10] (RSES) and KNIME.

In KNIME we have built workflow, which applied number of algorithms to our dataset by at first by binning selected attributes of the input data into the buckets. Size and type of buckets where specific to each model we built and is noted in the results

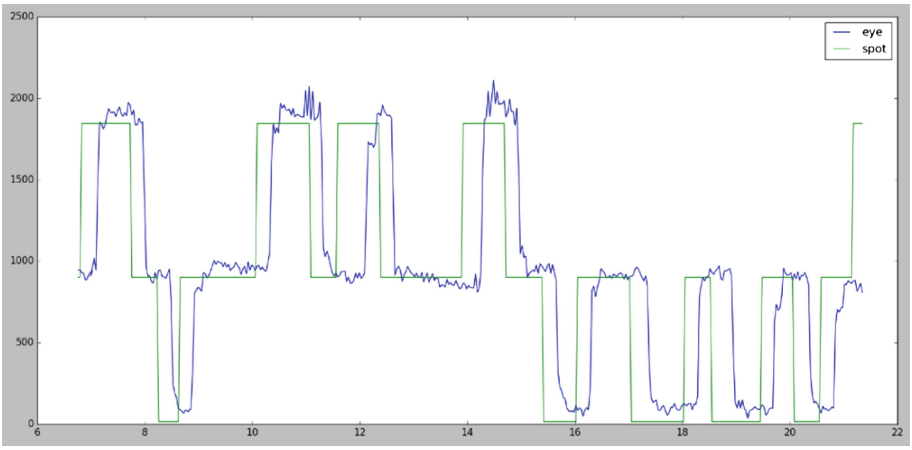

Fig. 1. This is the plot of eye position in time domain (x-axis). This recording was from the ET system. On the y-axis are position of eye gaze (ragged) and light spot (squared line)

section per classifier. KNIME provide binning methods for numeric values using two algorithms: equal bin and frequency bin method, first bin values based on minimum and maximum values to achieve numerically equal subsets while frequency bin try to perform binning so each subset have equal number of elements.

In RSES binning were implemented using built in discretization function [10], which included generating local cuts for number attributes excluding symbolic attributes.

Independently from used method, we have applied n-fold cross validation function, number of folds were specific for given model and as well as exact process parameters is noted in the results section.

3 Results

In the first part we built classifiers for predicting total UPDRS. As an input table for our model we have used data partly shown in Table 1.

We have built model for predicting total UPDRS. In order to verify predictions of our model (rules) we have used n-fold validation by dividing our set into n groups in two situations: (1) with- and (2) without-results of the eye movement measurements. In this way, it was possible to find importance of the reflexive saccades (RS) measurements on our predictions. We have binned total UPDRS into 4 groups with equal frequency algorithm, as described in the methods section. We have used RSES applying discretization with local cuts (excluding symbolic attributes) and the decision tree algorithm with 3-fold cross validation. Accuracy reached without RS data was 38.9%.

Next, we have added to our dataset RS attributes based on ET solution and applied similar process as for data in the first step. Table 2 presents information table after discretization.

Table 1. An example of data table for building classifier for total UPDRS in RSES. Legend for rows: mtre: mean delay right eye, stre: standard deviation for delay in right eye, stdredur/stdledur: standard deviation for saccade duration in right/left eye, latency_mean: mean latency for both eyes.

Patient #	'13/PD/BMT/2013'	'14/PD/POP/2010'	'55/PD/DBS/2013'	'56/PD/DBS/2013'
YearOfBirth	1948	1979	1955	1948
Sex	0	0	0	1
MonthsAfterDBS	18	48	12	12
Weight	61	58	70	88
BMT_dosage	750	400	500	MISSING
UPDRS_I	3	0	3	0
UPDRS_II	20	18	8	7
UPDRS_III	30	53	8	8
UPDRS_IV	2	2	3	2
UPDRS_TOTAL	55	73	22	17
Hoeh & Yahr scale	2.5	3	1	1
SchwabEnglandScale	70	60	90	90
PDQ39	77	34	26	49
AIMS	0	0	11	0
Epworth	9	7	6	12
ET_latency_mean	0.23	0.26	0.24	0.26
ET_latency_sd	0.06	0.09	0.08	0.08
ET_maxspeed_mean	1.73	0.41	0.53	0.97
ET_dur_mean	0.54	0.39	0.44	0.43
ET_dur_sd	0.54	1.04	0.43	0.48
Session	'S1'	'S1'	'S4'	'S4'

As states in Table 2 discretization algorithm selected only few significant attributes in order to create classification rules. Those included 5 RS attributes not only proving importance of RS saccades but also showing that ET parameters can be efficient in building classifier for PD. In contrast to model described in classification of dataset

Table 2. Discretized table for building classifier for predicting total UPDRS using ET RS data. Legend for RS attributes as in Table 1.

Patient #	'13/PD/BMT/2013'	'14/PD/POP/2010'	'14/PD/POP/2010'	'14/PD/POP/2010'
YearOfBirth	(−Inf, 1971.5)	(1971.5, Inf)	(1971.5, Inf)	(1971.5, Inf)
PDQ39	(46.5, Inf)	(30.5, 46.5)	(30.5, 46.5)	(30.5, 46.5)
AIMS	(−Inf, 5.5)	(−Inf, 5.5)	(−Inf, 5.5)	(−Inf, 5.5)
ET__mtre	(−Inf, 0.27)	(−Inf, 0.27)	(−Inf, 0.27)	(0.27, Inf)
ET__stre	(−Inf, 0.055)	(0.055, Inf)	(−Inf, 0.055)	(0.055, Inf)
ET__stdredur	(0.3349, Inf)	(−Inf, 0.3349)	(−Inf, 0.3349)	(0.3349, Inf)
ET__stdledur	(0.255, Inf)	(0.255, Inf)	(0.255, Inf)	(0.255, Inf)
ET__latency_mean	(−Inf, 0.255)	(0.255, Inf)	(−Inf, 0.255)	(0.255, Inf)
Session	'S1'	'S1'	'S2'	'S3'
UPDRS_TOTAL	(39, 64]	(64, 95]	(22, 39]	(39, 64]

without RS data we were able to achieve 72.2% accuracy, which is significantly better result.

We have performed another test using RS data from Ober saccadometer, applying the same method. The discretized table is shown below as Table 3.

Table 3. Discretized table with RS from Ober.

Patient #	'13/PD/BMT/2013'	'14/PD/POP/2010'	'14/PD/POP/2010'	'14/PD/POP/2010'
Sc_LatencyMeanLEFT	(−Inf, 217.0)	(263.0, 330.0)	(−Inf, 217.0)	(263.0, 330.0)
Sc_AmplitudeRIGHT	(10.35, 12.1)	(−Inf, 10.35)	(12.10, Inf)	(−Inf, 10.35)
Sc_PeakVelocityLEFT	(335.0, Inf)	(335.0, Inf)	(335.0, Inf)	(335.0, Inf)
Sc_LatencyMeanALL	(−Inf, 305.0)	(−Inf, 305.0)	(−Inf, 305.0)	(−Inf, 305.0)
Sc_DurationALL	(48.5, Inf)	(48.5, Inf)	(48.5, Inf)	(−Inf, 48.5)
session	'S1'	'S1'	'S2'	'S3'
UPDRS_TOTAL	(39, 64]	(64, 95]	(22, 39]	(39, 64]

As shown in Table 3 attributes significant for building classifier in our model again relay on RS, also we can note that some of the attributes like latency are used in both Ober and ET approach. Accuracy for this model was similar to model based on RS using ET and reached 66.7%.

Next we built model for predicting session number for PD patients. We have followed similar path as in building classifier for total UPDRS. Results for different datasets are shown in Table 4.

Table 4. Result for predicting session number using different classifiers.

	Accuracy
No RS data	58.30%
ET RS data	85%
Ober RS data	82.20%

There are similar trends as in case of prediction of total UPDRS, dataset with no RS had lowest accuracy. Those with RS data either from Ober or ET have similar accuracies; in case of prediction of the session number we are getting even better outcomes than for total UPDRS. Highest accuracy we have reached using combined dataset including RS from both ET and Ober systems.

We have combined list of attributes used in classification after discretization process of each dataset to show significant attributes used in building our models, those are shown in Table 5.

In the last step of our study we have run other common ML algorithms using KNIME on the datasets that proved to give best results in RSES. Additionally we have calculated other measures such as Cohen's kappa and Matthews Correlation Coefficient (MCC). As we can see in Table 5 we are able to reach high accuracy for predicting total UPDRS using standard algorithms reaching 85.7% and 71.4% respectively for WEKA decision tree algorithm and random forest (Table 6).

Table 5. Table showing significant attributes used to build model for predicting session number. Asterisk mark which columns were used while running classifier on given dataset.

	No RS data	ET RS data	OB RS data
Patient #	*	*	*
YearOfBirth	*	*	*
MonthsAfterDBS	*	*	*
UPDRS III	*	*	*
UPDRS IV	*		
ET_latency_sd		*	
OB_duration_mean			*
OB_amplitude_mean			*

Table 6. Results for building classifiers for predicting total UPDRS using common ML algorithms, including measures as Cohen's kappa and Matthews correlation coefficient.

	Accuracy	Cohen's Cappa	MCC
WEKA - decision table	85.71%	0.808	0.818
Random forest	71.43%	0.618	0.611
WEKA - random forest	57.14%	0.434	0.499
Tree ensemble	61.90%	0.488	0.421

4 Conclusions

Our study proved that we have successfully used a low cost ET eye tracker for clinically relevant eye movement measurements. We have measured parameters of the reflexive saccades (RS) and with help of the discretization process choose only relevant ones. We have built models for predicting PD patient session number as different treatments effectiveness and total UPDRS as general patient conditions. Our predictions of session number and UPDRS had a high accuracy when ET was used as well as the commercial saccadometer was utilized.

We have confirmed that fast eye movements are important biomarker for PD. Both devices, clinically approved Ober saccadometer and adapted by us to the clinical measurements - Eye Tribe can be used to improve diagnostic of PD symptoms.

Bibliography

1. Tsanas, A., Little, M.A., McSharry, P.E., Spielman, J., Ramig, L.O.: Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. IEEE Trans. Biomed. Eng. **59**, 1264–1271 (2012)
2. Salvatore, C., Cerasa, A., Castiglioni, I., Gallivanone, F., Augimeri, A., Lopez, M., Arabia, G., Morelli, M., Gilardi, M.C., Quattrone, A.: Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and progressive supranuclear palsy. J. Neurosci. Methods **222**, 230–237 (2014)

3. Prashanth, R., Dutta Roy, S., Mandal, P.K., Ghosh, S.: High-accuracy detection of early Parkinson's disease through multimodal features and machine learning. Int. J. Med. Inf. **90**, 13–21 (2016)
4. 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, Heidelberg (2014). doi:10.1007/978-3-319-09891-3_32
5. Szymanski, A., Szlufik, S., Dutkiewicz, J., Koziorowski, D.M., Cacko, M., Nieniecki, M., Przybyszewski, A.W.: Data mining using SPECT can predict neurological symptom development in Parkinson's patients. In: 2015 IEEE 2nd International Conference on Cybernetics (CYBCONF), pp. 218–223 (2015)
6. Przybyszewski, A.W., Kon, M., Szlufik, S., Dutkiewicz, J., Habela, P., Koziorowski, D.M.: Data mining and machine learning on the basis from reflexive eye movements can predict symptom development in individual Parkinson's patients. In: Gelbukh, A., Espinoza, F.C., Galicia-Haro, S.N. (eds.) MICAI 2014. LNCS (LNAI), vol. 8857, pp. 499–509. Springer, Heidelberg (2014). doi:10.1007/978-3-319-13650-9_43
7. Przybyszewski, A.W., Kon, M., Szlufik, S., Szymanski, A., Habela, P., Koziorowski, D.M.: Multimodal learning and intelligent prediction of symptom development in individual Parkinson's patients. Sensors **16**, 1498 (2016)
8. Pawlak, Z.: Rough Sets. Springer, Dordrecht (1991)
9. Bazan, J.G., Nguyen, H.S., Nguyen, S.H., Synak, P., Wróblewski, J.: Rough set algorithms in classification problem. In: Polkowski, P.L., Tsumoto, P.S., Lin, P.T.Y. (eds.) Rough Set Methods and Applications, pp. 49–88. Physica-Verlag HD, Heidelberg (2000)
10. Bazan, J.G., Szczuka, M.: RSES and RSESlib - A Collection of Tools for Rough Set Computations. In: Ziarko, W., Yao, Y. (eds.) RSCTC 2000. LNCS (LNAI), vol. 2005, pp. 106–113. Springer, Heidelberg (2001). doi:10.1007/3-540-45554-X_12