2018 17th IEEE International Conference on Machine Learning and Applications

Detailed Analysis of the Luria's Alternating series Tests for Parkinson's Disease Diagnostics

Sven Nõmm
Konstantin Bardõš
Department of Software Science,
Faculty of Information Technology
TalTech,
Akadeemia tee 15 a,
12618, Tallinn, Estonia
E-mail: sven.nomm@ttu.ee.

Aaro Toomela
School of Natural Sciences
and Health,
Tallinn University
Narva mnt. 25, 10120,
Tallinn, Estonia
E-mail: aaro.toomela@tlu.ee

Kadri Medijainen
Institute of Sport Sciences
and Physiotherapy
University of Tartu
Puusepa 8, Tartu 51014, Estonia
E-mail: kadri.medijainen@ut.ee

Pille Taba
Department of Neurology
and Neurosurgery
University of Tartu
Puusepa 8, Tartu 51014, Estonia
E-mail: pille.taba@kliinikum.ee

***Abstract*—Patterns drawn by the patients during digital Luria's alternating series tests are analysed in this paper to support diagnostics of the Parkinson's disease. There are two main components that distinguish the approach proposed in this paper. The first one is the application of digital Luria's alternating series tests. In spite of its simplicity battery of the Luria's alternating series, tests allow not only to diagnose the disease but also may help to uncover on which level motor functions are affected. The second component is that kinematic and geometric features are computed not on the basis of the entire pattern but its "logical" constituents. Such an approach is justified by the fact that there is no formal description of the errors possible during the testing. Finally, it is demonstrated that classification decisions may be traced and interpreted.**

I. Introduction

Analysis of the patterns drawn during digital Lurias alternating series tests (dLAST) with respect to aid Parkinsons disease (PD) diagnostics constitutes research subject of the present paper. PD is a progressive disorder [1] characterised by many symptoms, including rigidity, bradykinesia, flexed posture, and freezing [2]. Consequently, the quality of life of patients with PD is markedly reduced [3], [4]. However, currently, the diagnosis of (PD) is still based on clinical criteria based on the presence of a combination of cardinal symptoms [5]. As manifestations of PD affect various aspects of life, the need for an early biomarker or diagnostic test is evident. This motivates the research directed towards developing computer-aided systems to support diagnostics of PD. Fine motor tests and especially drawing and writing tests were used to diagnose PD and other neurodegenerative diseases for a long period of time. Usually conducted using paper and pen, tests were assessed by practitioners. In such setting time and the number of observed errors are the only countable parameters of the test. Introduction of digital tables and later tablet computers has allowed to register and analyse kinematic parameters of the tests and pressure applied by the stylus pen on the screen. At the same time, error detection lacks behind human practitioners. At this point, it is worth to note that there is no formal description of what error is (obviously for each test such set of errors and their descriptions may vary). The present research aims to tackle two main problems. The first one is to build a classifier to support diagnostics of the PD on the basis of dLAST. The second one is to provide the basis for the formal description of errors particular to this type of testing. dLAST tests were not studied as extensively as spiral drawing or clock drawing tests and may be regarded as novel component brought by the present research. Recently digitisation of existing test results was discussed in [6] which demonstrates the importance of the tests. Another novel component is that set of features computed to describe kinematics of the drawing process. Finally, the authors made an attempt to trace (interpret) classification process using LIME technique proposed in [7]. The paper is organised in the form of five sections. Background information and formal problem statement are presented in Section II. Presentation of the experimental setting and methodology chosen for the current research constitutes Section III. Section IV discusses achieved results. Finally, conclusions are drawn in the last section.

II. Problem statement and background information

Formally, the problem studied within the frameworks of this research belong to the area of supervised learning and may be stated as follows: based on the labelled set of drawn Luria's alternating series extract (generate) and select the subset of features possessing discriminating power to distinguish between

978-1-5386-6805-4/18/$31.00 ©2018 IEEE
DOI 10.1109/ICMLA.2018.00219

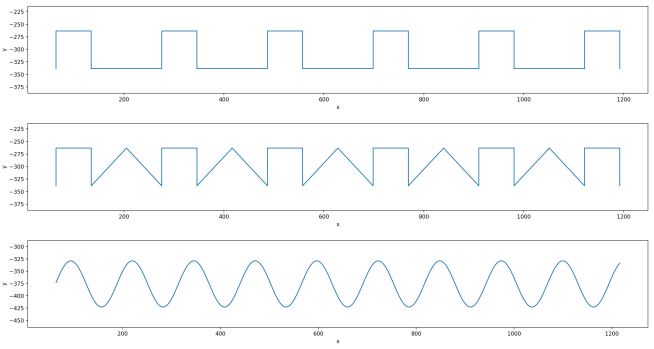

Fig. 1. Reference paths shown to the tested individual

PD patients and healthy control (HC) individuals of the same age. Using this subset of features train the classifiers to aid diagnostics procedure. For the training purposes, it is essential to have a group of the individuals with diagnosed PD. This will provide ground truth information to assign labels to the test results. Drawing and writing tests have been used in psychiatry and neurology for many decades. Introduction of digital tables and tablet computers has sparked research activities directed to digitise the tests. In [8] just four parameters were recorded and analysed. Recent contributions are usually base their results on a much larger number of extracted features, for example, [9], which makes features selection procedure necessary. There are many published results devoted to the computer-aided diagnostics and modeling of PD. Initially proposed in [10] and [11] alternating series test is targeted to assess the state of planning and execution functions of fine motor motions. During the test individual is asked to continue, copy and trace certain periodic path referred as series. Initially, three types of series were suggested. Π-type and $\Pi\Lambda$-type and Sinusoidal line, depicted in Figure 1 The test requiring to continue the series is targeted on assessment of the state of planning function. The test requiring to trace the line is targeted to assess the state of implementation function. For the test requiring to continue the line only part of the series should be shown. Digitisation of the Luria's alternating series tests was initially proposed by [12], where it was demonstrated that some of kinematic and pressure parameters are distinguishable between the groups of PD patients and healthy control (HC) individuals of the similar age. Contrary to [12] the present research suggest to divide drawn paths into the straight segments and describe each segment by the set of kinematic, geometric, temporal and pressure parameters. Then proper feature selection procedure is applied. Finally, the classification algorithm is trained and evaluated. Within the present contribution, only Π-path and $\Pi\Lambda$-path is considered to keep conscience of the analysis part, whereas numeric results are presented for $\Pi\Lambda$ series only. Corresponding results for Π series are available from the author upon request.

III. Experimental setting and methodology

A. Experimental setting

Experimental setting used in the present research consists of three main components. Apple iPad Pro 9.7-inch tablet with Apple Pencil equipped with the specially developed application was used to conduct the tests. The application was developed by the members of the research team and allows to choose the battery of tests, record anonymous code of the tested individual (in accordance to the personal data protection law and permit of the ethics committee), conduct the tests and send acquired data to the remote server for storage and processing. Once the battery of tests is chosen, the person conducting the testing positions the tablet in front of the tested individual on the table. Ensures that right hand of the tested person is supported by the surface of the table and provides them with Apple Pencil (stylus). Remaining part of the process is no different compared to the case where paper and pen is used. During the test, the application collects the information about the position of the stylus tip in the coordinate system of the screen. Also, the pressure applied by the stylus tip and stylus orientation may be recorded. This information is recorded up to 200 times per second and supplied by the corresponding time stamp. As a result, each test is described by $n \times 6$ matrix where n is the number of time instances observed during the test. Once testing is complete this information is transferred to the remote server in the form of Jason files. Remaining part may be conducted either immediately on the server or on any reasonably powerful computer. Π-path and $\Pi\Lambda$-path is considered to keep conscience of the analysis part.

B. Data acquisition

Dataset used within the present research were acquired by the professional physiotherapist. It covers 15 HC individuals and 15 PD patients (individuals with diagnosed PD) of approximately the same age, (mean value 65) with equal gender proportions.

C. Data processing

The data processing workflow is organising in three consequent stages. The first one is the preprocessing stage where the data was filtered to eliminate noise occurring during unintentional contacts of the pen and screen. Then velocity and acceleration are computed for each observation point. Assessment performed by a human is based on evaluating corners between consequent straight segments of the drawn series here and after referred as *edges*.Formally the edge may be defined as a line-segment, represented by vector of points $[p_i, p_{i+1}, ...p_{j-1}, p_j]$ where $[p_i, p_j]$ are starting and ending points of the segment. Therefore, it is essential to detect joint points of the consequent edges. On the second stage, Shi-Tomasi algorithm [13] is applied to find the segments corresponding to the edges of the reference drawing. It is the classical algorithm used in machine vision and therefore utilises only geometric properties of the drawing. Drawn series and its segmentation into the edges are depicted in figure 2. On the final stage of data processing kinematic, geometric

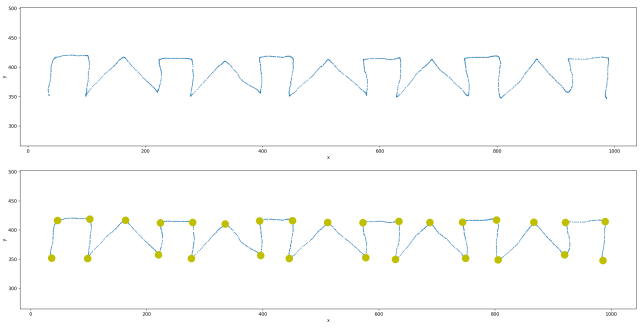

Fig. 2. Drawn ΠΛ series and its segmentation into the edges.

temporal and pressure based features are computed for each detected edge. As a result of data processing, each test is described by a numeric vector of features.

D. Feature extraction

The features computed for each edge. are assigned to three different classes: edge features, drawing features, kinematic and pressure features. Drawing features describe drawing in general and on the first view should not be considered in the research concentrating its attention on the analysis of individual segments of the drawing. At the same time drawing features were suggested in [12] therefore evaluating these features allows demonstrating advantages of the current results over those reported in [12]. Edge features are defined with respect to each individual edge. Table I presents the subset of edge features.

TABLE I
SAMPLE SUBSET OF EDGE FEATURES

Edge Feature	Description
angle	Angle of the line between points $[p_i, p_j]$
distance	Euclidean distance between points $[p_i, p_j]$
duration	Time interval between points $[p_i, p_j]$
speed	Linear speed between points $[p_i, p_j]$
pressure_median	Median pressure in vector of points $[p_i, p_{i+1} \cdots, p_{j-1}, p_j]$
pressure_max	Max pressure in vector of points
pressure_min	Min pressure in vector of points
longitude_median	Median longitude of a stylus in vector of points
longitude_max	Max longitude of a stylus in vector of points
longitude_min	Min longitude of a stylus in vector of points
latitude_median	Median latitude of a stylus in vector of points
latitude_max	Max latitude of a stylus in vector of points
latitude_min	Min latitude of a stylus in vector of points

Kinematic and pressure feature set is composed on the basis of sets proposed by [14], [15], [12] and have proven high level of discrimination power between PD patients and HC individuals. Kinematic and pressure based feature set describes an arbitrary vector of points $[p_1, p_2, ...p_{n-1}, p_n]$, therefore can be defined and evaluated in the scope of both *Drawing* and *Edge* entities. Along with standard *[median, mean, mass]* features, another higher-order feature was proposed - *number of changes*. The subset of kinematic and pressure features is presented in Table II. The naming of the features is analogous to the case of edge features.

TABLE II
SAMPLE SUBSET OF KINEMATIC AND PRESSURE FEATURES

Feature	Description
duration	Time period between first and last point of vector $[p_1, p_2, ...p_{n-1}, p_n]$
trajectory_length	Sum of all Euclidean distances between all neighbour points in vector
velocity_mass	Velocity mass of the point vector $[p_1, p_2, ...p_{n-1}, p_n]$
velocity_mean	Average velocity of the point vector
velocity_nc	Number of velocity changes in point vector
acceleration_mass	Acceleration mass of the point vector
acceleration_mean	Average acceleration of the point vector
acceleration_nc	Number of acceleration changes in point vector
jerk_mass	Jerk mass of the point vector
jerk_mean	Average jerk of the point vector
jerk_nc	Number of changes in jerk of point vector
pressure_diff_mean	Average difference in pressure between neighbour points
pressure_mass	Pressure mass of the point vector
pressure_nc	Number of changes in pressure

Finally, the subset of drawing features is presented in Table III these features are computed for the entire Luria's alternating pattern.

TABLE III
SAMPLE SUBSET OF DRAWING FEATURES

Drawing Feature	Description
width	Horizontal distance of the *Drawing* $x_{max} - x_{min}$
height	Vertical distance of the *Drawing* $y_{max} - y_{min}$
area	*Drawing height* multiplied by *width*
number_of_strokes	Number of strokes. Corresponds to maximum stroke *index* in vector of points $[p_1, p_2, ...p_{n-1}, p_n]$
number_of_edges	Number of *Edge* elements in the Drawing, calculated from number of *Edges* in graph on depth level $d = 0$
completeness	Ratio between *actual* and *expected* number of *Edge* elements in the *Drawing* pattern. Expected number of *Edges* is Luria pattern-type related *constant*
angle_up._regression_line	Angle of *upper* regression line
angle_low._regression_line	Angle of *lower* regression line
angle_mid._regression_line	Angle of *middle* regression line
angle_up._low.	Angle between *upper* and *lower* regression lines of the *Drawing*

E. Feature selection

Feature selection consisted of two stages executing filter based selection procedures. On the first stage, statistical hypothesis testing was used to select the features which differ significantly between PD patients and HC individuals. Two-sample t-test to reject *null hypothesis* and determine if datasets differ significantly from each other. In the current context, *null hypothesis* is that the samples representing HC and PD subjects belong to the independent populations with *equal* means. *Alternative hypothesis* samples representing HC and

PD subjects originate from the populations with *unequal* means, hence there is the significant difference between data sets. As a result of this stage, a subset of features that differ significantly between two populations is found. On the second stage Fisher's score (1) [16] is used to order selected features with respect to their discriminating power.

$$F = \frac{\sum_{j=1}^{k} p_j(\mu_j - \mu)^2}{\sum_{j=1}^{k} p_j \sigma_j^2} \tag{1}$$

F. Classifiers

For the purpose of best performing classifier acquisition for PD discrimination task, it was decided to train, analyse and compare classical and modern machine learning algorithms from the following subset:

- k-NN k-nearest neighbours ($k = 3$)
- Decision Tree
- Random Forest
- Support Vector Machine (SVM)
- Ada Boost

Trained classifiers were validated using K-Fold cross validation. Classifier training and validation was performed using Python programming language and *Scikit-learn* library.

IV. Main results

A. Feature selection

As a result of two stage feature selection described in Subsection III-E the fallowing subsets of the features were selected. In Table IV kinematic and edge features with highest Fisher's scores are presented together with corresponding results of t-test Totally 78% of $n = 872$ features from

TABLE IV
FEATURE SELECTION RESULTS: EDGE FEATURES

Feature Name	*Fisher* score	p-value	t-stat
edge_03_speed	0.9000	6.11×10^{-10}	8.60
edge_15_velocity_mean	0.7802	1.12×10^{-9}	8.38
edge_13_speed	0.7649	1.51×10^{-11}	10.03
edge_03_velocity_mean	0.7591	6.70×10^{-10}	8.57
edge_15_acceleration_mean	0.7464	4.87×10^{-9}	7.84
edge_15_jerk_mean	0.7442	3.40×10^{-8}	7.15
edge_03_acceleration_mean	0.6709	2.06×10^{-9}	8.15
edge_11_velocity_mean	0.6662	2.00×10^{-10}	9.02
edge_14_speed	0.6050	4.35×10^{-10}	8.73
edge_11_acceleration_mean	0.5998	5.14×10^{-10}	8.67

current cluster passed t-test with significance level $\alpha = 0.05$. Highest observed Fisher score $F = 0.90$ received feature *edge_03_speed*, i.e, linear speed of third edge of the drawing, which means, that speed within particular segment of the pattern provides much higher discrimination potential, than any kinematic feature of full drawing ($F = 0.50$ was observed for *drawing_acceleration_mean*). Selection of the drawing features with highest Fisher's scores is presented in Table V Comparing Tables IV and V one may easily see that edge features are more informative for discriminating between the PD patients and HC individuals.

TABLE V
Drawing FEATURES — STATISTICAL ANALYSIS

Feature Name	*Fisher* score	p-value	t-stat
drawing_acceleration_mean	0.5025	2.56×10^{-11}	9.82
drawing_jerk_mean	0.4415	3.42×10^{-10}	8.82
drawing_velocity_mean	0.4071	1.67×10^{-11}	9.99
drawing_slope_mass	0.3555	1.37×10^{-13}	12.01
drawing_acceleration_nc	0.3372	6.29×10^{-6}	5.36
drawing_velocity_nc	0.3333	1.46×10^{-5}	5.08
drawing_jerk_nc	0.3313	3.80×10^{-6}	5.54
drawing_pressure_nc	0.3261	3.96×10^{-9}	7.92
drawing_duration	0.3149	2.26×10^{-8}	7.29
anomalies_drawing_length	0.2021	3.68×10^{-8}	7.12

B. Classifiers training and evaluation

To analyse different combinations of models and hyperparameters and achieve high classification accuracy P_{acc}, collection of models with *k-NN, Decision Tree, Random Forest, SVM and AdaBoost* was trained with different number of top-performing features (n = 3, $n = 30$, $n = 90$) from separate *Drawing, Edge, Anomaly* and *Mixed* features-clusters for each of $k = 3$ folds. All randomly generated folds with trained models and accuracy measures were saved in auxiliary *Classifier-container entity* for subsequent analysis. In Table VI evaluation results are presented for the classifiers trained on the subset of three best features. In Table VII evaluation results

TABLE VI
CLASSIFIER ACCURACY P_{acc} — TRAINED WITH TOP $n = 3$ FEATURES

Classifier	P_{acc} all	P_{acc} drawing	P_{acc} edge
AdaBoost	0,79	0,79	0,76
DecisionTree	0,82	0,76	0,79
k-NN	0,71	0,84	0,71
RandomForest	0,86	0,84	0,88
SVM$_{linear}$	0,56	0,50	0,67
SVM$_{rbf}$	0,56	0,62	0,56

are presented for the classifiers trained on the subset of thirty best features.

TABLE VII
CLASSIFIER ACCURACY P_{acc} — TRAINED WITH TOP $n = 30$ FEATURES

Classifier	P_{acc} all	P_{acc} drawing	P_{acc} edge
AdaBoost	0,77	0,70	0,80
DecisionTree	0,70	0,80	0,80
k-NN	0,74	0,69	0,70
RandomForest	0,86	0,77	0,91
SVM$_{linear}$	0,53	0,50	0,50
SVM$_{rbf}$	0,56	0,56	0,62

In Table VIII evaluation results are presented for the classifiers trained on the subset of thirty best features. From classifier type perspective — one can observe, that best performing model is *Random Forest* with highest accuracy $P_{acc} = 0.91$, high performance also demonstrates *AdaBoost* and classical *Decision Tree* with $P_{acc} = 0.86$ and $P_{acc} = 0.88$. For some reason, Support Vector Machine classifies the dataset with

TABLE VIII
CLASSIFIER ACCURACY P_{acc} — TRAINED WITH TOP $n = 90$ FEATURES

Classifier	P_{acc} all	P_{acc} drawing	P_{acc} edge
AdaBoost	0,82	0,77	0,86
DecisionTree	0,79	0,69	0,88
k-NN	0,65	0,71	0,71
RandomForest	0,85	0,80	0,89
SVM$_{linear}$	0,50	0,50	0,49
SVM$_{rbf}$	0,56	0,62	0,56

slightly lower accuracy rate of $P_{acc} = 0.67$ From feature space perspective one can achieve $P_{acc} = 0.88$ with only $n = 3$ features. Which conforms to high statistical significance of extracted drawing parameters.Best accuracy of $P_{acc} = 0.91$ was obtained with $n = 30$ feature space. Much higher dimensionality of $n = 90$ features does not boost model accuracy, but even makes it slightly lower with $P_{acc} = 0.88$. From feature class perspective — *Edge*-related features perform best with highest accuracy $P_{acc} = 0.91$. Fusion of all features yields classifier with $P_{acc} = 0.86$. *Drawing*-related features describe Luria pattern in general and can produce model with $P_{acc} = 0.84$.

C. Decision tracing

Classification results based on 3-feature models may be relatively easy to interpret. For example by positioning a particular point in 3D graph with drawn decision boundary and scatter plot of training points. At the same time, models based on 90 are rather black boxes and may be unacceptable for the delicate area of clinical applications. Local Interpretable Model-Agnostic Explanations (LIME) algorithm proposed by [7] is meant to overcome this gap. LIME essentially an explanation technique, that explains the predictions of *any* classifier in an interpretable and faithful manner, by learning an interpretable model locally around the prediction. It is assumed that any complex classification model is locally linear and therefore linear model can be used to explain prediction s of the complex model locally. Examples of decision tracing with LIME algorithm is presented in Table IX for the case of PD patient and in Table X for the HC individual

On the one hand the threshold values for each feature, presented in the inequalities of the first column require provide the foundation to develop formal description of the errors which may be observed by practitioner. But on the other hand it is still necessary analyse these values to provide better interpretation in terms familiar to medical community.

V. DISCUSSION

Up to this point only so-called ΠΛ series were considered. The same methodology was applied to the case of Π series. Obtained results did not possess any difference and feature values and training results are available from the authors upon request. It should be noted that exactly the same procedure is applied to the cases of continue, copy and trace test types. If a particular individual is classified as healthy by the trace test and as PD by the continue trace type this is a clear indication that

TABLE IX
PD INSTANCE EXPLANATION EXAMPLE

Classifier Model — RandomForest, $n = 90$ features		
Instance PD-07 — *Prediction* PD (71%) / Control (29%)		
Prediction Explanations: Parkinson's Disease PD (71%)		
Feature Explanation	Feature Actual	Significance
edge_03_velocity_mean $\leq$ 9.07e+03	4.25e+03	0.061013
edge_14_pressure_nc $>$ 1.95e+01	2.80e+01	0.054542
edge_03_acceleration $\leq$ 1.92e+08	7.28e+07	0.042965
drawing_slope_mass $>$ 1.83e+05	3.21e+05	0.037471
1.65e+01 $<$ edge_06_pr.. $\leq$ 1.90e+01	1.90e+01	0.034442
edge_03_jerk_mean $\leq$ 1.21e+13	3.43e+12	0.029171
edge_22_speed $\leq$ 5.39e+01	3.90e+01	0.027016
1.62e+13 $<$ edge_00_jerk $\leq$ 2.76e+13	2.14e+13	0.018656
1.25e+00 $<$ edge_00_pr.. $\leq$ 1.51e+00	1.45e+00	0.012146
5.32e+01 $<$ edge_05_sp.. $\leq$ 7.01e+01	5.85e+01	0.011786
6.27e+01 $<$ edge_12_sp.. $\leq$ 1.02e+02	1.00e+02	0.011575
edge_14_duration $>$ 9.40e-01	1.62e+00	0.010385
Prediction Explanations: Healthy Control HC (29%)		
Feature Explanation	Feature Actual	Significance
edge_14_velocity_mean $\leq$ 7.31e+03	5.81e+03	-0.012261
edge_14_acceleration.. $\leq$ 1.43e+08	1.36e+08	-0.012769
2.26e+08 $<$ edge_07_a... $\leq$ 3.20e+08	2.29e+08	-0.013645
5.81e+03 $<$ edge_10_v... $\leq$ 9.51e+03	7.19e+03	-0.015295
6.81e+12 $<$ edge_20_j... $\leq$ 1.32e+13	7.37e+12	-0.023397
6.23e+01 $<$ edge_10_sp.. $\leq$ 1.16e+02	9.00e+01	-0.025807
5.96e+03 $<$ edge_15_v... $\leq$ 1.20e+04	9.29e+03	-0.032646
7.28e+01 $<$ edge_06_sp.. $\leq$ 1.05e+02	7.65e+01	-0.077155

TABLE X
HEALTHY CONTROL HC INSTANCE EXPLANATION — EXAMPLE

Classifier Model — RandomForest, $n = 90$ features		
Instance HC-01 — *Prediction* HC (90%) / PD (10%)		
Prediction Explanations: Healthy Control HC (90%)		
Feature Explanation	Feature Actual	Significance
edge_06_speed $>$ 1.05e+02	2.67e+02	0.111788
edge_11_velocity_mean $>$ 1.45e+04	2.63e+04	0.074923
edge_15_velocity_mean $>$ 1.82e+04	2.23e+04	0.073401
drawing_slope_mass $\leq$ 1.14e+05	8.27e+04	0.059881
edge_03_velocity_mean $>$ 1.56e+04	3.79e+04	0.057738
edge_06_pressure_nc $\leq$ 1.05e+01	5.00e+00	0.057620
edge_14_pressure_nc $\leq$ 1.02e+01	7.00e+00	0.037617
edge_03_jerk_mean $>$ 2.73e+13	7.69e+13	0.032639
edge_10_speed $>$ 1.16e+02	2.60e+02	0.031768
edge_22_speed $>$ 1.21e+02	2.67e+02	0.030603
edge_21_acceleration $\leq$ 6.75e+00	4.00e+00	0.026311
edge_05_speed $>$ 9.92e+01	2.46e+02	0.024375
edge_10_velocity_mean $>$ 1.50e+04	3.28e+04	0.023728
edge_03_acceleration $>$ 3.54e+08	8.96e+08	0.021507
edge_14_duration $\leq$ 4.20e-01	3.30e-01	0.018153
Prediction Explanations: Parkinson's Disease PD (10%)		
Feature Explanation	Feature Actual	Significance
edge_00_jerk_mean $>$ 2.76e+13	3.54e+13	-0.021943
edge_01_jerk_mean $>$ 1.70e+13	4.17e+13	-0.026075
edge_14_velocity $>$ 1.29e+04	2.30e+04	-0.027546
edge_14_acceleration $>$ 2.92e+08	6.07e+08	-0.037125
edge_20_jerk_mean $>$ 2.73e+13	7.93e+13	-0.066437

disease has affected the planning function but implementation function remains unaffected.

VI. Conclusions

The main result of the present research has clearly demonstrated the possibility to use digital Luria's alternating series tests to support diagnostics process of the Parkinson's disease. It was shown that segmentation of the series into the edges lead significantly more informative feature set and in turn provides rich ground to construct accurate classifiers. Finally, a novel technique was applied to trace classification decisions. On the one hand it provides the possibility to interpret classification results in a more human friendly form. On the other hand it gives foundation to formally describe errors observed during the testing process. Future research will be directed towards adopting this technique to the sinusoidal series.

References

[1] W. Poewe, K. Seppi, C. M. Tanner, G. M. Halliday, P. Brundin, J. Volkmann, A.-E. Schrag, and A. E. Lang, "Parkinson disease," *Nature reviews Disease primers*, vol. 3, p. 17013, 2017.

[2] J. Jankovic, "Parkinsons disease: clinical features and diagnosis," *Journal of neurology, neurosurgery & psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.

[3] E. D. Louis and D. G. Machado, "Tremor-related quality of life: a comparison of essential tremor vs. parkinson's disease patients," *Parkinsonism & related disorders*, vol. 21, no. 7, pp. 729–735, 2015.

[4] A. Schrag, M. Jahanshahi, and N. Quinn, "What contributes to quality of life in patients with parkinson's disease?" *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 69, no. 3, pp. 308–312, 2000.

[5] A. Tarakad and J. Jankovic, "Diagnosis and management of parkinson's disease," in *Seminars in neurology*, vol. 37, no. 02. Thieme Medical Publishers, 2017, pp. 118–126.

[6] P. Stpien, J. Kawa, D. Wieczorek, M. Dabrowska, J. Slawek, and E. Sitek, *Computer Aided Feature Extraction in the Paper Version of Lurias Alternating Series Test in Progressive Supranuclear Palsy*, P. Ewa, B. Pawel, K. Jacek, and W. Wojciech, Eds. Springer, 2018.

[7] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you?: Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 1135–1144.

[8] C. Marquardt and N. Mai, "A computational procedure for movement analysis in handwriting," *Journal of Neuroscience Methods*, vol. 52, no. 1, pp. 39 – 45, 1994. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0165027094900531

[9] P. Drotar, J. Mekyska, I. Rektorova, L. Masarova, Z. Smekal, and M. Faundez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of parkinson's disease," *Artificial Intelligence in Medicine*, vol. 67, pp. 39 – 46, 2016.

[10] A. R. Luria, *Higher Cortical Functions in Man*. Springer, 1995.

[11] J. R. Hodges, *Cognitive Assessment for Clinicians Second Edition*. Oxford Medicine, 2007.

[12] S. Nomm, A. Toomela, J. Kozhenkina, and T. Toomsoo, "Quantitative analysis in the digital luria's alternating series tests," in *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Nov 2016, pp. 1–6.

[13] J. Shi *et al.*, "Good features to track," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*. IEEE, 1994, pp. 593–600.

[14] P. Zham, D. K. Kumar, P. Dabnichki, S. Poosapadi Arjunan, and S. Raghav, "Distinguishing different stages of parkinsons disease using composite index of speed and pen-pressure of sketching a spiral," *Frontiers in Neurology*, vol. 8, p. 435, 2017.

[15] M. San Luciano, C. Wang, R. A. Ortega, Q. Yu, S. Boschung, J. Soto-Valencia, S. B. Bressman, R. B. Lipton, S. Pullman, and R. Saunders-Pullman, "Digitized spiral drawing: A possible biomarker for early parkinsons disease," *PloS one*, vol. 11, no. 10, p. e0162799, 2016.

[16] C. C. Aggarwal, *Data Mining: The Textbook*. Springer Publishing Company, Incorporated, 2015.