2018 15th International Conference on
Control, Automation, Robotics and Vision (ICARCV)
Singapore, November 18-21, 2018

Interpretable Quantitative Description of the Digital Clock Drawing Test for Parkinson's Disease Modelling

Sven Nõmm, Ilja Masharov, Aaro Toomela, Kadri Medijainen and Pille Taba

Abstract— **Modelling of the fine motor motions during digital clock drawing test is performed within frameworks of the present studies to facilitate computer-aided diagnostics of the Parkinsons disease. Clock drawing test has been used to diagnose and monitor neurodegenerative diseases for a long period of time. It was one of the first tests to be digitalized. Nevertheless, its natural complexity causes many problems for the analysis of test results. Unlike simpler tests clock test drawing consists of many different elements. Therefore, it requires one to identify all the elements of the drawing and then proceed with the analysis of different elements. The presence of different elements, in turn, leads the idea to model test results on different levels. Low-level where modelling is performed on the basis of kinematic, pressure and temporal parameters describing fine motor movements. And higher level, where relative positions of elements, drawing quality and presence of the different elements are analyzed. Low-level analysis constitutes the scope of the present paper. Based on the clock drawing test results obtained for the groups of patients with diagnosed Parkinsons disease and similar, by age and size, healthy individuals. Features describing fine motor motions are constructed, whereas special attention is paid to the so-called set of motion mass parameters. Then features possessing the highest discriminative power to distinguish between Parkinsons disease patients and healthy control individuals are selected. Finally based on the selected subset of features applicability of different machine learning algorithms to support diagnostics process is evaluated.**

I. Introduction

Applicability of the kinematic parameters, recorded during the digital clock drawing test, for diagnostics of the Parkinson's disease (PD) constitutes the scope of the present research. PD belongs to the most widely spread of neurodegenerative diseases. Developing PD usually affects motor functions of the patient. Most common symptoms of the disease are the rigidity, unintentional movements and tremor [1]. Needless to say, that such symptoms may seriously complicate the everyday life of the patient. While there is no known cure for the PD being diagnosed in the early stages and properly treated, the symptoms may be reduced[2]. Therefore ability to diagnose PD on the early stages has a crucial importance for the quality of everyday life of the patient. Unfortunately, diagnostics of PD on early stages is not an easy task. This motivates the idea to provide enhanced tools to support diagnosing process. During the recent years number of solid results has demonstrated that gross- and fine- motor functions analysis may be used for computer-aided diagnostics of the neurodegenerative diseases [3], [4]. Clock drawing test belongs to the area of fine motor testing and widely used in neurology [5], [6]. It is usually performed by means of the paper and pen [7]. The tested individual is asked to draw the clock on a white sheet of paper. Practitioner observes visually the drawing process and then evaluates the final drawing. In its non-digitalized version, the main attention is paid to the positioning and sizes of the digits, quality of the ring and ability of the patient to correctly position clock hands [8]. Such setting has two important drawbacks. The first one is the subjectivity of human opinion. Even for highly trained practitioners assessments may vary [9]. The second one is the inability of the human eye to capture many parameters describing the drawing process. For example, velocities, accelerations and pressure applied by the pen tip to the screen could not be recorded. For more than thirty years attempts to digitise fine-motor tests has been made. This has resulted in a wide variety of features used to describe stylus movements [10], [11], [4]. Also, many attempts have been made to demonstrate that machine learning classifiers are the viable tools and may support diagnostics and modelling of the neurodegenerative diseases[12], [13]. Nevertheless, the medical community remains sceptical and such computer-supported disease diagnostics and modelling remains on the stage of testing. Such scepticism may be caused by two factors related to the interpretation of the results. Firstly decisions made by machine learning algorithms are not always traceable. Secondly, novel technologies provide a lot of new features but the meaning of those features may be difficult to interpret. The present research has concentrated its attention on these issues. It is demonstrated that the motion mass parameters describing kinematics of fine motor motions allow distinguishing between PD patients and healthy controls (HC). At the same time, these parameters are easy to interpret.

Unlike the simpler fine motor tests, clock drawing test requires one to produce a drawing with many different elements. This causes more complex analysis procedure to be applied. In the first stage, it is necessary to identify different elements of the drawing and only then analyse them separately. Formal problem statement is given in Section II. Background information and formal statement are given in SectionII. Experimental setting and applied techniques are

Sven Nõmm and Ilja Masharov are with Department of Software Science, School of Information Technology Tallinn University of Technology, Akadeemia tee 15 a, 12618, Tallinn, Estonia sven.nomm@ttu.ee.

Aaro Toomela is with School of Natural Sciences and Health, Tallinn University, Narva mnt. 25, 10120, Tallinn, Estonia aaro.toomela@tlu.ee

Kadri Medijainen is with Institute of Sport Sciences and Physiotherapy, University of Tartu, Puusepa 8, Tartu 51014, Estonia kadri.medijainen@ut.ee

Pille Taba is with Department of Neurology and Neurosurgery, University of Tartu Puusepa 8, Tartu 51014, Estonia pille.taba@kliinikum.ee

discussed in detail in Section III. Section IV presents main results of the present studies. Discussion of the achieved results and concluding remarks constitute the last section.

II. Problem statement and background information.

There is a large variety of published results confirming, that PD affects human motor system [14], [11]. In the early stages of the disease, this may be reflected by shaking, rigidity and slowness of motions. This should be reflected by the set of kinematic parameters describing motions of the patient [3]. On the fine motor level motions are usually captured by means of tablet computers or digital tables. These devices provide additional functionality to record information about stylus orientation and the pressure it applies to the screen of the device. This leads the working hypothesis consisting of two research questions.

1) There is a set of parameters (kinematic, static and temporal) describing clock drawing test, such that their values differ significantly between the groups of PD patients and similar in age and gender distribution group of healthy control (HC) individuals.
2) Based on these parameters it is possible to train classifies to diagnose Parkinson's disease on the basis of the clock drawing test.

III. Experimental Setting and Methodology

In order to answer the research questions stated in the previous section, one has to follow the standard work-flow of machine learning.

A. Data acquisition and hardware

The digitalisation of the clock drawing test was implemented among the battery of tests in the form of specialised application for Apple's iPad pro with 9.7-inch screen and Apple pen stylus. The application records the coordinates of the pen tip together with its orientation and pressure applied to the screen with the sampling rate of 200 frames per second and stores this information together with the corresponding time stamp in the file. Upon completion of the test, this file is transferred to the server for processing. The drawing itself may be represented as the set of dots depicted in 1

Digital clock drawing test was conducted for the groups of PD patients and a group of healthy controls with approximately the same age and gender distribution. Totally 30 individuals were tested. Note, that the data acquisition and analysis were conducted with proper permission granted by the ethics committee (detailed information is available on behalf of the authors upon request)).

B. Methodology

The feature extraction consists of three parts: extraction of clock circle, clock hands extraction and digits extraction. While the last one was done using third-party libraries available for python and Tensor Flow library the first two were implemented by the authors of the present paper. Drawing of the complex objects usually consists of many

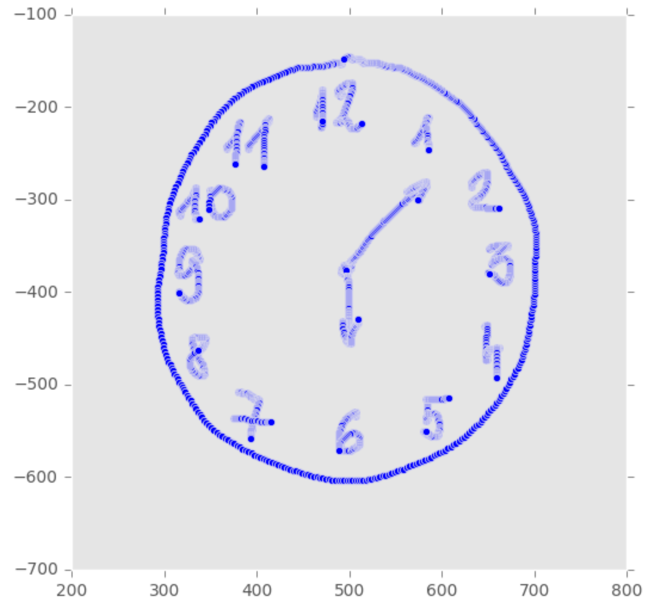

Fig. 1. Result of a clock drawing depicted as the set of points

strokes. Stroke may be seen as the primitive indivisible element of the drawing. The stroke is defined as the sequence of consequently recorded points, while the stylus stayed in constant contact with the surface. Some recent results have demonstrated the importance of the stroke analysis in digitisation of fine motor tests [10]. From the machine learning viewpoint, the following procedure may be treated as the stroke classification problem. Where each stroke should be assigned a label of the clock element it belongs to.

1) Extraction and analysis of the clock circle: Extraction of the clock circle is based on the following assumptions:

- The clock circle is drawn in a single stroke.
- The clock circle is the longest stroke.
- The clock circle is drawn first.
- The clock circle is the outermost stroke (closest to the boundaries of the drawing area).

The strongest of this assumption is the first one. If it does not satisfy, a number of features describing the process may not be computed. These assumptions are treated as heuristics whereas each assigned its own weight. Each stroke is tested against each heuristic. Based on the heuristic weights the probability of stroke to be the clock circle is computed. For the present research, the threshold was chosen experimentally and set to 0.8. If clock circle is not detected on the basis of heuristics then multi-stroke detection procedure may be applied after all other elements of the clock are detected. If clock circle is identified then two sets of features may be extracted. The first set are the geometric properties of the drawn circle and the second on are the kinematic, temporal and pressure parameters describing the drawing process. The first centre of the drawn clock circle is computed as the arithmetic mean of all the stroke points then three so-called "perfect circles" are calculated:

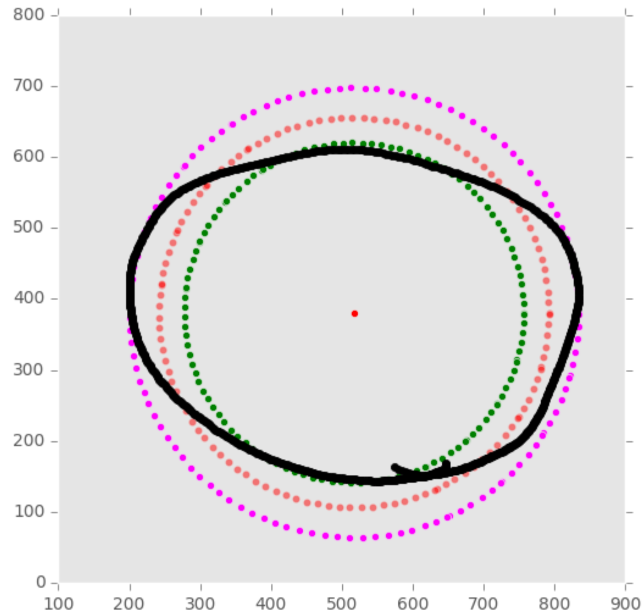

Fig. 2. Drawn clock circle together with the corresponding inner, outer and mean circles

- Inner circle: is the maximum inscribed circle into the drawn clock circle.
- Outer circle: is the minimum circumscribed circle.
- Mean circle is the circle which radius computed as the mean of the distances from the centre to each point of the drawn circle.

Drawn clock circle together with inner, outer and mean circles are depicted in Fig.2 Intuitively one may think that closer the drawing to the "perfect" circle, less chance of motor function disruptions. In addition to these measures [15] proposes two more "roundness" measures. Least square circle (LSC) formally defined as the circle that separates a drawn circle by separating the sum of total areas inside and outside of it in equal amounts. In this case, the error is the sum of the differences d_i between the maximum and minimum distances from the reference circle.

$$F(x_c, y_c, r_c) = \min(\sum_{i=1}^{n} d_i^2) \tag{1}$$

where (x_c, y_c) are the coordinates of the drawn clock circle and r_c is its radius. The second measure is minimum zone circle (MZC) is defined as the difference between the radiuses of outer and inner "perfect" circles.

2) Clock hands and digits detection: Clock circle recognition is followed by the clock hands detection and analysis. It is assumed that clock hands are drawn with two strokes closest to the circle centre. To identify digits of the clock deep convolutional neural network proposed by [16] was used. It was trained on the MIST data set [17]. Recognised elements of one particular test example are depicted in Figure 3 At this point, each individual stroke of the drawing may be classified as the part of a particular element of the drawing (digit, clock hand or circle) or treated as the noise. Noise elements are usually caused by the fact that some elements are missing (forgotten by the patient) or be simply unrecognizable. Also, unintentional contacts of the stylus with the surface may be the cause of noise elements.

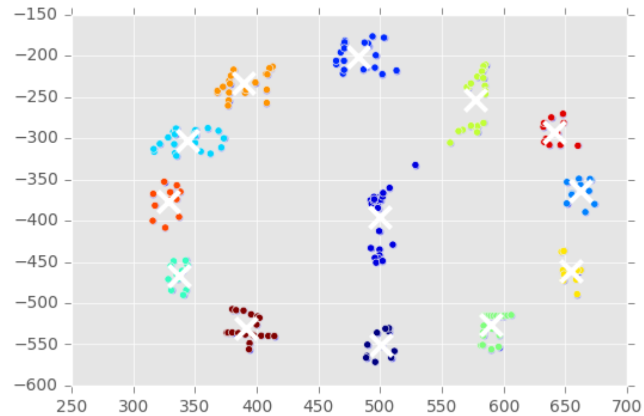

Fig. 3. Recognised elements of the clock drawing

C. Kinematic, pressure and temporal parameters describing the drawing

For each stroke of the drawing four types of attributes are extracted:

1) Geometric attributes: length, width, area of the minimal rectangle containing the stroke and number of points. Also features describing the goodness of the clock circle belongs to this category.
2) Temporal features: stroke drawing duration, pause duration between the strokes.
3) Kinematic features: velocity, acceleration, jerk and their directional versions.
4) Pressure: the pressure applied by the stylus tip to the screen of the device.

Based on these features motion mass parameters are computed. Initially proposed in [18] and [19] to describe smoothness and amount of the gross motor motions, necessary to perform a certain action or physical exercise. Later in [4] these parameters were adopted for the fine motor case, to describe the smoothness and the amount of the movements of the stylus tip. Unlike the values of descriptive statistics motion mass parameters are closer to the integral or motion energy description. For the sake of self-sufficiency let us briefly explain the main idea of the motion mass parameters. Remind, that each stroke (line) of the drawing is nothing else but the sequence of points. At each point, one may compute absolute values of velocity and acceleration along the tangent vector. Then for each stroke (or each element of the drawing) these values are summed up leading the values refereed as *velocity mass* (2) and *acceleration mass* (3).

$$V_M = \sum_{i=1}^{n} |v_i| \tag{2}$$

where v_i is the velocity at time instance i calculated along the tangent vector.

$$A_M = \sum_{i=1}^{n} |a_i|. \tag{3}$$

where a_i is the velocity at time instance i calculated along the tangent vector. Recently [10] has indicated the importance of jerk (acceleration of acceleration or third derivative of velocity) as an important parameter allowing to distinguish PD patients from HC group, *jerk mass* is computed in the analogous way to the velocity and acceleration masses.

$$J_M = \sum_{i=1}^{n} |j_i|. \tag{4}$$

where j_i is the velocity at time instance i calculated along the tangent vector. Motion mass features are complemented by two parameters describing total pressure change *pressure mass*

$$P_M = \sum_{i=1}^{n} |\rho_i|. \tag{5}$$

where ρ_i is the difference between the pressure measured at time instance i and pressure measured at time instance $i-1$. *angular mass* total direction changes observed for the given element.

$$D_M = \sum_{i=1}^{n} |d_i|. \tag{6}$$

where d_i is the angle between the tangent vectors in points i and $i-1$. The tuple $M_M = \{V_M, A_M, J_M, P_M, D_M\}$ is refereed as motion mass parameters. In its original form trajectory length, euclidean distance E_M between the beginning and ending points of the motion, time and ratios A_m/E_M and V_M/E_M were included in the tuple. Then the set of features, describing digital clock drawing test is computed as following. For each element of the drawing measures of descriptive statistics are computed attribute wise. Also the same measures are computed for all strokes observed during the test. (without splitting it into different elements) and complimented by overall time of the drawing. Feature extraction process is depicted in Figure4.This process results in 373 numeric features available for the analysis.

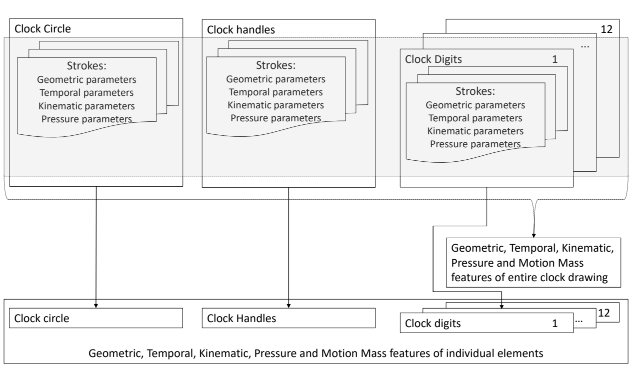

Fig. 4. Feature extraction process

D. Feature selection and model building

Two step filter selection technique is used for feature selection. On the first step statistical hypothesis testing is used to demonstrate existence of the feature subset that differ significantly between the groups of PD patients and healthy controls. Formally, for each feature the following pair of hypothesis was tested. The the statement of null (base) hypothesis $H_0 : \mu_{PD} = \mu_{HC}$ whereas the statement of alternative hypothesis $H_1 : \mu_{PD} \neq \mu_{HC}$. Since the samples describing PD patients and healthy controls may have an unequal variances Welch's test was chosen to perform the testing. Answering first research question this step also allows to demonstrate that proposed hard- and soft- ware setting is sensitive enough to distinguish between the patients and healthy controls. On the second step Fisher's score [20] is used to select the features with the highest discriminating power. This two-step feature selection procedure is followed by the application of four machine learning techniques: decision tree classifiers, k -nearest neighbours, logistic regression and support vector machines.

IV. Main results

Welch's test result has demonstrated that among 373 numeric features, describing the clock drawing test following 20 differ significantly between the groups of PD patients and healthy controls.

TABLE I
FEATURES THAT DIFFER SIGNIFICANTLY BETWEEN PD AND HC GROUPS

Feature	t statistic	p-value	Fisher's score
Strokes: average V_M	-4.1415	0.0009	0.7146
Total drawing duration	2.3331	0.0287	0.2267
Circle angular mass	2.7392	0.015	0.3126
Hands average J_M	-2.4493	0.027	0.2499
Strokes std. Acc.	-5.0088	7.7681e-05	1.0453
Strokes average J_M	-4.3618	0.0007	0.7927
Circle average V_M	-3.9402	0.0009	0.6469
Strokes std. slope	-2.2078	0.0385	0.2031
Strokes std. J_M	-5.1125	6.8307e-05	1.0890
Longest stroke duration	3.2215	0.0052	0.4324
Circle average A_M	-4.0943	0.0006	0.6984
Hands average A_M	-2.2236	0.0413	0.2060
Circle average J_M	-3.9900	0.0007	0.6633
Circle drawing duration	3.6629	0.002	0.5591
Strokes std. V_M	-4.0235	0.0005	0.6745
Strokes average A_M	-4.1845	0.0009	0.7295
Circle average slope	-2.1755	0.0404	0.1972
Drawing totlal duration	4.2473	0.0003	0.7516
Strokes average D_M	-2.9307	0.0076	0.3578
Strokes average J_M	-2.3157	0.034	0.2234

Based on the Fisher's score values following features were selected for classifier training: Standard deviation of the jerks masses computed on the basis of all observed strokes, average jerk mass of all strokes and standard deviation of the acceleration masses of all observed strokes. The number of features allowed for classifier training is limited by the total sample size of thirty individuals. Scatter plot representing the distribution of HC individuals (blue color) and PF patients (red color) is presented in Figure 5. Performance of the four most popular shallow-learning classifiers: decision

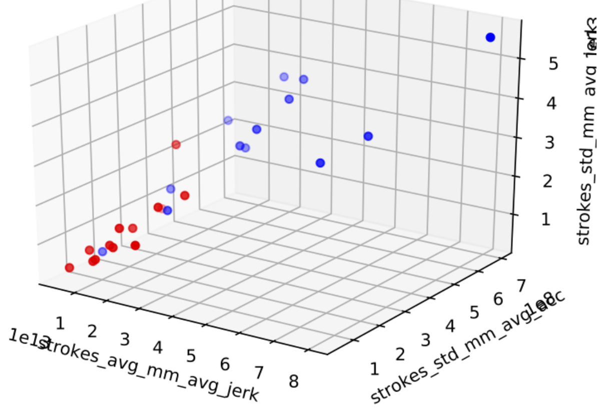

Fig. 5. Scatter plot of depicting distribution of PD patients and healthy controls described by three features with the highest fisher scores.

trees (DT), k - nearest neighbors (kNN), logistic regression (LR) and support vector machine (SVM) were evaluated. 6-fold cross-validation procedure was used to evaluate the performance of the classifiers. Average accuracy score for the decision tree classifier is 0.72 and for the k - nearest neighbours classifier (k = 3) 0.83. SVM and logistic regression classifiers has demonstrated accuracy scores slightly above 0.5. Maximal accuracy score fore the decision tree classifier is 0.81, corresponding confusion matrix is presented in Table II Maximal accuracy score fore the kNN classifier is 0.9,

TABLE II
CONFUSION MATRIX - DECISION TREE CLASSIFIER

	Actual (PD)	Actual (HC)
Predicted (PD)	5	1
Predicted (HC)	1	4

corresponding confusion matrix is presented in Table III. Many scholars consider that temporal features are the most

TABLE III
CONFUSION MATRIX - kNN CLASSIFIER

	Actual (PD)	Actual (HC)
Predicted (PD)	5	1
Predicted (HC)	0	5

informative for diagnosing and modeling of PD, an attempt was made to replace one of the chosen features with drawing total duration and drawing duration of the longest stroke. As a result, average accuracy has decreased in both cases. While kNN has demonstrated a slightly higher accuracy level, the difference is mostly caused by false positive numbers in the case of decision tree classifier. Also, decision tree classifier may be represented by the tree-like graph depicted in Figure 6. This graph may be easily interpreted as the sequence of conditions verify.

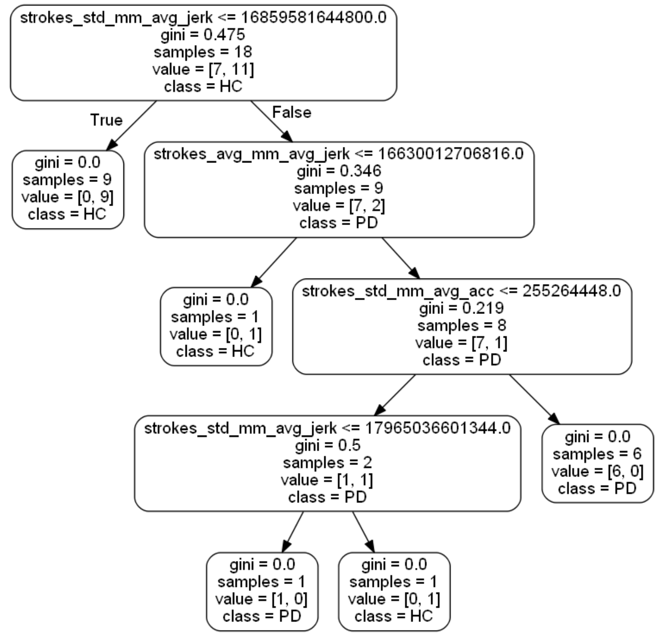

Fig. 6. Decision tree classifier.

V. CONCLUSIONS

The present paper proposes the method to evaluate clock drawing test results by means of features describing the amount and smoothness of the stylus tip. First statistical hypothesis testing has demonstrated that proposed hard- and soft- ware setting is sensitive enough to distinguish between the Parkinson's disease patients and similar healthy control individuals. Then features possessing the highest discriminating power were selected. These features are based on the accelerations and jerks of the stylus tip movements and therefore describe the smoothness of the drawing process. Evaluation of the four simple machine learning classifiers has demonstrated that decision trees and k - nearest neighbors classifiers are the most suitable to support the diagnostics process. Also, it was demonstrated, that while temporal parameters differ significantly between the PD patients and controls they are less informative compared to the kinematics parameters.

REFERENCES

[1] J. Jankovic, "Parkinsons disease: clinical features and diagnosis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.

[2] L. V. Kalia and A. E. Lang, "Parkinson's disease," *The Lancet*, vol. 386, no. 9996, pp. 896 – 912, 2015.

[3] S. Nõmm, A. Toomela, M. Vaske, D. Uvarov, and P.Taba, "An alternative approach to distinguish movements of parkinson disease patients," in *The 13th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*, August 2016, pp. 49–19.

[4] 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.

[5] K. A and K. A, "On drawing impairment in alzheimer's disease," *Archives of Neurology*, vol. 48, no. 1, pp. 73–77, 1991. [Online]. Available: + http://dx.doi.org/10.1001/archneur.1991.00530130083024

[6] K. Yurinosuke, U. Hideki, N. Jin, N. Kaeko, K. Hitoshi, and F. Kenji, "Qualitative analyses of clock drawings in alzheimer's disease and vascular dementia," *Psychiatry and Clinical Neurosciences*, vol. 55, no. 5, pp. 485–491, 2001.

[7] J. J. d. Paula, D. M. d. Miranda, E. N. d. Moraes, and L. F. Malloy-Diniz, "Mapping the clockworks: what does the Clock Drawing Test assess in normal and pathological aging?" *Arquivos de Neuro-Psiquiatria*, vol. 71, pp. 763 – 768, 10 2013.

[8] M. C. Lessig, J. M. Scanlan, H. Nazemi, and S. Borson, "Time that tells: critical clock-drawing errors for dementia screening," *International Psychogeriatrics*, vol. 20, no. 3, pp. 459–470, 2008.

[9] W. Souillard-Mandar, R. Davis, C. Rudin, R. Au, and D. L. Penney, "Interpretable machine learning models for the digital clock drawing test," *CoRR*, vol. abs/1606.07163, 2016.

[10] 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.

[11] S. Joella, R. Jeffrey, B. David, and C. David, "A comparison of five clock scoring methods using roc (receiver operating characteristic) curve analysis," *International Journal of Geriatric Psychiatry*, vol. 16, no. 4, pp. 394–399, 2001.

[12] S. Nõmm, K. Bardõš, I. Mašarov, J. Kozhenkina, A. Toomela, and T. Toomsoo, "Recognition and analysis of the contours drawn during the poppelreuter's test," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec 2016, pp. 170–175.

[13] P. Drotar, J. Mekyska, Z. Smekal, I. Rektorova, L. Masarova, and M. Faundez-Zanuy, "Prediction potential of different handwriting tasks for diagnosis of parkinson's," in *E-Health and Bioengineering Conference (EHB)*, Nov 2013, pp. 1–4.

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

[15] W. Sui and D. Zhang, "Four methods for roundness evaluation," *Physics Procedia*, vol. 24, pp. 2159 – 2164, 2012, international Conference on Applied Physics and Industrial Engineering 2012.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, ser. NIPS'12. USA: Curran Associates Inc., 2012, pp. 1097–1105. [Online]. Available: http://dl.acm.org/citation.cfm?id=2999134.2999257

[17] Y. LeCun and C. Cortes, "MNIST handwritten digit database," 2010. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[18] S. Nõmm and A. Toomela, "An alternative approach to measure quantity and smoothness of the human limb motions," *Estonian Journal of Engineering*, vol. 19, no. 4, p. 298308, 2013.

[19] S. Nõmm, A. Toomela, and J. Borushko, "Alternative approach to model changes of human motor functions," in *European Modelling Symposium EMS2013 (EMS2013)*, Manchester, United Kingdom, Nov. 2013.

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