TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Vassili Gorbatšov 193522 IAIB

MOTION FREEZING ANALYSIS IN DRAWING AND WRITING TESTS

Bachelor's Thesis

Supervisor: Sven Nõmm
PhD

Tallinn 2024

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Vassili Gorbatšov 193522 IAIB

LIIKUMISPEATUSTE ANALÜÜS KIRJUTAMISE JA JOONISTAMISE TESTIDES

Bakalaureusetöö

Juhendaja: Sven Nõmm
PhD

Tallinn 2024

Author's Declaration of Originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Vassili Gorbatšov

27.05.2024

Abstract

The current study focuses on the phenomenon of upper limb freezing in Parkinson's disease patients during writing tests. With technology development appeared new possibilities of the disease examination and diagnosis. Digitisation of the writing and drawing tests coupled with artificial intelligence techniques have proven effective in Parkinson's disease detection. However the analysis of freezing episodes did not get much attention. Previous studies were mainly focused on the analysis of freezing episodes which occurred during lower limbs movements.

The primary objective of the present thesis is to determine whether the neighbourhood of the point where freezing occurred is enough to effectively differentiate between Parkinson's disease patients and healthy control subjects. For each freezing episode are considered time intervals of half seconds before, after freezing and a whole time frame around freezing. These intervals are described by the hand movement's kinematic and pressure parameters, which are used as features for the standard machine learning workflow that applies a nested cross-validation loop.

The key findings of the study show that analysing the areas around freezing episodes can help to distinguish Parkinson's disease patients from healthy individuals. The best results were obtained from movements occurring half second post-freezing. Usage of movements kinematic and pressure-based features in training classifiers have helped to reach values of 0.86, 0.86 and 0.93 for accuracy, precision, and recall. Achieved results are comparable to those available in the literature.

The thesis is written in English and is 30 pages long, including 8 chapters, 8 figures and 14 tables.

Annotatsioon
Liikumispeatuste analüüs kirjutamise ja joonistamise testides

Käesolev lõputöö keskendub ülemiste jäsemete peatumise nähtusele, mis esineb Parkinsoni tõvega patsientidel lausetestide kirjutamise ajal. Tehnoloogia arenguga tekkisid uued võimalused haiguse uurimiseks ja diagnoosimiseks. Kirjutamis- ja joonistamistestide digitaliseerimine koos tehisintellekti tehnikatega on osutunud tõhusaks meetodiks Parkinsoni tõve tuvastamiseks. Tardumisepisoodide analüüsi peale pole suurt tähelepanu enne pööranud. Eelmised uuringud keskendusid peamiselt alajäsemete liigutuste ajal esinevate tardumisepisoodide uurimisele.

Käesoleva lõputöö põhieesmärk on teha kindlaks, kas tardumise toimumise koha naabruspiirkond on piisav, et kindlalt eristada Parkinsoni tõvega patsiente ja terveid kontrollisikuid. Iga tardumis episoodi puhul vaadeldakse ajavahemiku poole sekundi ulatuses enne, pärast tardumist ja kogu ajaintervalli tardumise ümber. Need ajavahemikud on kirjeldatud käe liikumise kineetiliste- ja surveparameetrite abil, mida kasutatakse standardse masinõppe töövoo tunnuseks, mille peale rakendatakse ristvalideerimise tsüklit.

Uuringu peamised tulemused näitavad, et peatumispiirkondade ümbritsevate alade analüüs võib aidata eristada Parkinsoni tõvega patsiente tervetest inimestest. Parimad tulemused saadi liikumiste puhul, mis toimusid pool sekundi jooksul pärast tardumist. Liigutuste kineetiliste ja survepõhiste tunnuste kasutamine klassifikaatorite treenimisel aitas saavutada väärtusi $0,86$, $0,86$ ja $0,93$ *accuracy*, *precision* ja *recall* puhul. Saavutatud tulemused on sarnased kirjanduses kättesaadavate tulemustega.

Lõputöö on kirjutatud Inglise keeles ning sisaldab teksti 30 leheküljel, 8 peatükki, 8 joonist, 14 tabelit.

List of Abbreviations and Terms

PD	Parkinson's Disease
HC	Healthy Control
ML	Machine Learning
LR	Logistic Regressions
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
DT	Decision Tree
RF	Random Forest
AB	AdaBoost
JSON	JavaScript Object Notation

Table of Contents

List of Figures

List of Tables

1. Introduction

Parkinson's disease (PD) is a neurodegenerative disorder characterised by unintentional movements, rigidity, tremors, and freezing [1]. PD symptoms can severely affect the patient's lifestyle and quality [2]. There is no known cure for PD, but appropriate treatment can alleviate symptoms that affect patient quality of life. Early diagnosis often leads to more effective treatment results; however, it is difficult to precisely diagnose PD [3], especially in its early stages, and distinguish it from other diseases.

The findings of [4] show that PD frequently affects fine motor movements. Recent studies demonstrate that writing and drawing tests can be used to support the diagnosis of PD [5], [6]. The digitisation of writing and drawing tests has begun with the seminal paper [7]. Although digital tables are the primary medium for data collection [8], some recent results demonstrate that the use of tablet PC devices is applicable [9], [10] for the same tasks.

In the area of sentence writing, the analysis is based on the kinematic and pressure description of the entire test or its elements: words and individual letters [11]. The present research turns its attention to the phenomenon of *freezing* observed during writing tests [12] which, along with the phenomenon of micrographia [13], [14] is considered important to support the diagnosis of PD.

The preliminary results of this research have been published in [15]. The present thesis concludes the results of research and extends them to analyse the impact of literacy on the state of motor functions. AI tools were used only for spell checking.

The thesis is organised as follows. Section 1 consists of the introduction and analysis of related research. Section 2 defines the problem and the main objectives. Section 3 provides an overview of the data used in the thesis and its collection. Section 4 describes data processing and analysis methodologies. Section 5 describes the results. Section 6 consists of the illiteracy analysis process and the results. Sections 7 and 8 offer a discussion and conclusions about the acquired results and future research strategies.

2. Problem statement

The analysis of digital drawing and writing tests has become a trend during the last decade, but very little attention has been paid to analysing freezing. There are two main reasons for this. The first one is that freezing itself is not properly defined, e.g. different medical schools have different definitions of freezing. In turn, it is not straightforward to detect it technically. Freezing can be confused with pauses between word writings or with sharp turns during the letter writing process.

The working hypothesis of the present research assumes that the kinematic and pressure parameters of the hand movements observed in certain time neighbourhood around the freezing episode allow us to distinguish if the PD patient or HC subject performs the test.

2.1 Main Objectives

The objectives of the thesis may be stated as follows:

- Recognise where the written sentence was affected by the hand/fingers freezing.
- Describe freezing parameters such as duration, position of the pen tip, and pressure on the tablet screen.
- Describe the movements in the immediate vicinity of freezing.
- Determine to what extent machine learning techniques can be used to diagnose PD and illiteracy based on freezing parameters.

3. Background

The present thesis is part of a larger research that studies human motor functions whose objectives are to support the diagnosis of neurodegenerative diseases, detect early cognitive impairments, and recognise signs of fatigue. The experimental setup for this research is described in [9], [11] and [10]. The main facts about the hardware and software setup and the testing process are listed in Sections 3.1 and 3.2.

The data acquisition process is carried out under strict privacy law guidelines. The study is approved by the Research Ethics Committee of the University of Tartu (No.$1275T - 9$).

3.1 Hardware and software

Data used in present thesis was acquired during previous research. For data gathering during writing tests, Apple iPad pro (2016) with a 9.7-inch screen and Apple pen (stylus) were used. To collect movements of the stylus tip, software and interface suitable for the task was developed. The coordinates of the apple pen tip and the pressure applied to the screen were saved to the matrix. The rows of the matrix correspond to the observation points acquired up to 200 times per second, and the columns contain information that describes each point.[9] For each test collected data were saved for future processing in JavaScript Object Notation (JSON) files.

An example of data in a JSON file is shown below in (Figure 1). For each point was recorded information about the X coordinate (*x*) and the Y coordinate (*y*), pressure applied to the screen (*p*), stylus orientation altitude (*l*) and azimuth (*a*), time stamp (*t*).


```
 "data":[[{"x":74.0625,"l":0.7661090000000004,"a":2.301463,"y":142.6523,"p":0.333332
"t":531589068.96154398},{"x":74.593800000000002,"l":0.7661090000000004,"a":2.301463,"
.1172,"p":0.33333299999999999,"t":531589069.02279902},{"x":75.328100000000006,"l":0.7
.7661090000000004,"a":2.301463,"y":140.78909999999999,"p":0.33333299999999999,"t":53
{"x":76.25,"l":0.7661090000000004,"a":2.301463,"y":140.5898,"p":0.33333299999999999,"
"t":531589069.023287},{"x":76.718800000000002,"l":0.7661090000000004,"a":2.301463,"y"
.53129999999999,"p":0.33333299999999999,"t":531589069.02339798},{"x":77.109399999999
.7661090000000004,"a":2.301463,"y":140.53129999999999,"p":0.33333299999999999,"t":53
.02355099},{"x":77.843800000000002,"l":0.7661090000000004,"a":2.301463,"y":141.7148,"
.33333299999999999,"t":531589069.023669},{"x":78.640600000000006,"l":0.766109000000000
.44139999999999,"p":1.616403,"t":531589069.023785},{"x":79.640600000000006,"l":0.7661
.7661090000000004,"a":2.301463,"y":149.15629999999999,"p":2.2139959999999999,"t":531
```


Figure 1. Part of data in json file.

3.2 Handwriting data acquisition

In the experiment took part group of volunteers consisting of 24 patients with Parkinson's disease and 30 healthy control subjects, whose native language is *Estonian*. In the result of the experiment was created a DraWritePD dataset. Each participant completed the battery with 12 tests. In this experiment two test types were used: drawing and writing. In the present thesis, only a sentence writing test is considered.

In the sentence writing test, volunteers were asked to hand write the sentence *Kui Arno isaga kooli-majja jõudsid, olid tunnid juba alanud*, which means *When Arno with his father arrived, school lessons had already started*. The sentence is taken from the book learnt in all the schools in Estonia and was familiar to all participants. Figure 2 represents the sentence written on the screen of a tablet.

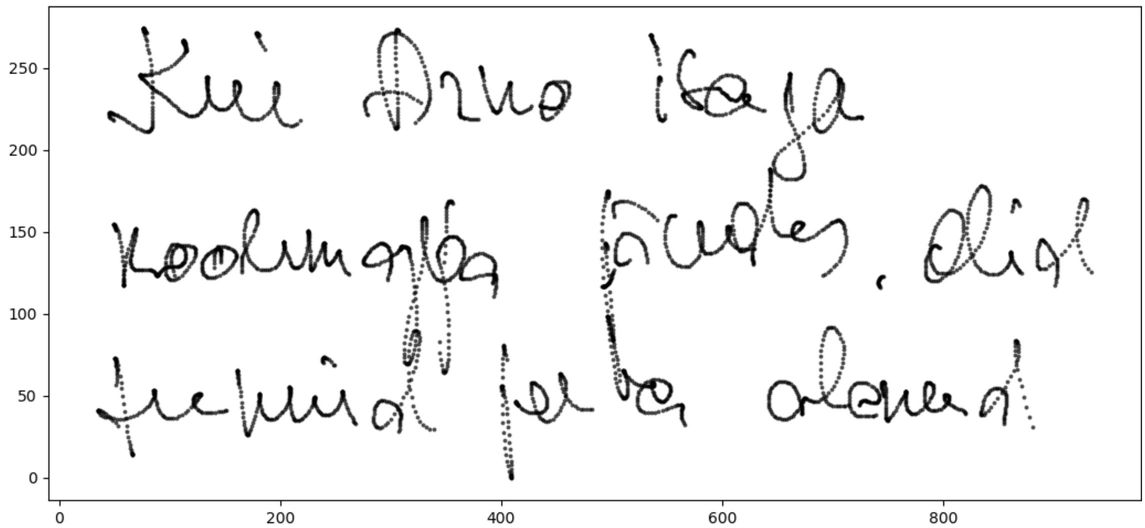

Figure 2. Sentence written on the screen of a tablet.

4. Methodology

Data processing was divided into several steps. Each of them was performed offline. The first step consists of freezing recognition. In the second step neighbourhood corresponding to 0.5 sec. around freezing is extracted. During the third step, machine learning workflow hosting the nested cross-validation loop for the six most frequently used machine learning classifiers: Logistic regression (LR), Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Decision Tree (DT), Random Forest (RF), and AdaBoost (AB) was applied. Nested cross-validation helps to prevent overoptimistic model goodness by avoiding bias and overfitting [16], [17].

Research, analysis, and tuning of hyper parameters were performed using Python programming language and PyCharm IDE. Following open-source Python libraries were extensively used on different study stages:

- NumPy and Pandas | for manipulation with data
- Matplotlib, Plotly and Seaborn | for figure and graph plotting
- Scikit-learn | for training and validation of numerous classifier models

4.1 Feature extraction

Feature extraction was done based on developed in [18] and [10] approach. The first step is similar to the results of [7] and [8]. It requires describing motion kinematic parameters by computing velocity, acceleration and higher position derivatives: jerk - J_N, yank Y_N, tug - T_N, snatch - Sn_N and shake - Sh_N [10]. Parameters describing the angle between the directional vectors in two consecutive observation points, integral-like parameters named *motion mass* [18], and descriptive statistics measures: maximum, mean, variance, supplement original feature set of parameters.

Equation 4.1 shows how these parameters are calculated in the example of velocity mass. In equation 4.1 N is the number of observation points, v_i the velocity (computed on the basis of two neighbouring points along the directional vector of the stylus tip movement) i where $i \in \{1, \ldots, N\}$.

$$V_N = \sum_{i=1}^{N} |v_i| \tag{4.1}$$

Calculation of other kinematic and pressure parameters uses same logic

4.2 Freezing episode

The freezing of the hand during the writing or *freezing episode* is defined as a *sudden, variable, and often unpredictable transient break in movement* [19]. To allow automatic detection of freezing [12] have adapted the definition to *an involuntary stop or clear absence of effective writing movements during at least* 1 *second*. The last approach could be implemented in the form of programming code. It should be considered that small jigging may occur during the hand freezing episode. For this research it was assumed, that freezing episode should have moment with total absence of movements. This moment was considered as middle point of the freezing. The timestamps of the beginning and ending points of freezing provide information about the extracted intervals. In this research, the length of this interval was experimentally found to be 1 of a second. Original freezing episode definition was split into three parts: Point with total absence of movement became the freezing and intervals of 0.5 seconds before and after the freezing became freezing neighbourhoods. Figure 3 shows freezing episodes, numbered and marked by yellow lines with green arrows.

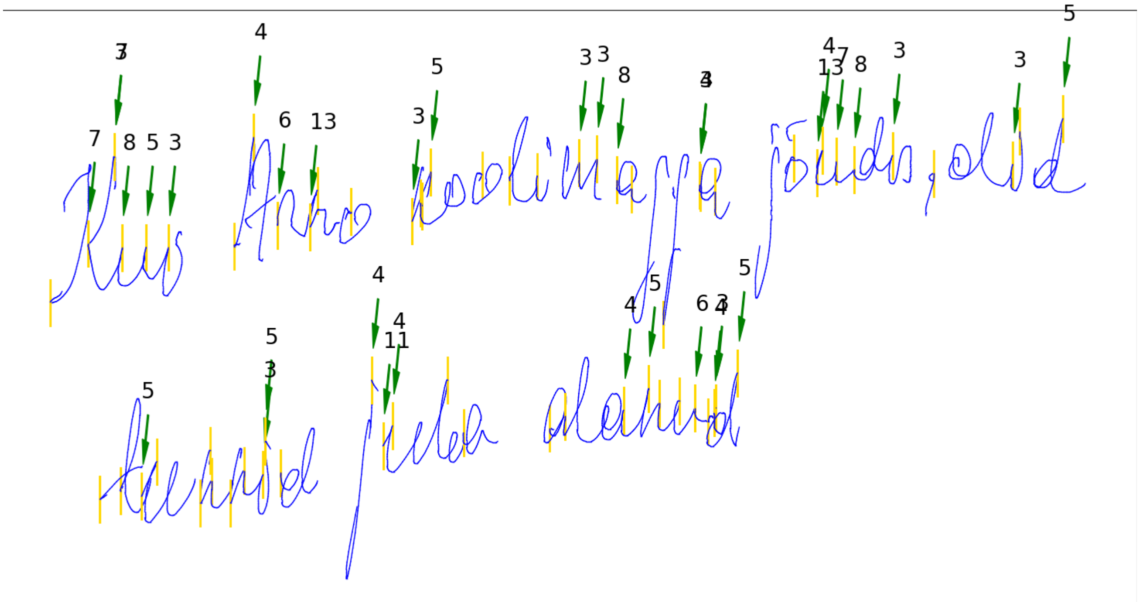

Figure 3. Freezing episodes.

The number of freezing episodes may vary between the PD and HC subjects test cases. It required to apply proper sampling to avoid problems caused by unbalanced data sets. After that feature extraction procedure was applied to these intervals. Dataset for ML analysis consisted of tuples with kinematic, pressure and motion mass parameters of freezings, what were calculated for each test. For every tuple was also added label marking that parameters belongs to PD patient or HC subject.

4.3 Feature selection and classification

Based on the size of the data set in this research was used filter models technique. Other feature selection techniques require larger data sets. Fisher's score 4.2 was chosen because the kinematic, pressure, and motion mass parameters of the movements of the tip of the stylus are numeric [20]. Fisher's score is a filter model feature selection technique, which assigns a numeric value to each feature, according to their discriminating power.

$$F = \frac{\sum_{i=1}^{N} p_j(\mu - \mu_i)^2}{\sum_{i=1}^{N} p_i \sigma_i^2} \tag{4.2}$$

In (4.2) N is the number of classes, μ and μ_i are the mean value of the entire set along the given characteristic and the mean value of the class i, respectively; p_i proportion of the class i and σ_i is the standard deviation of the class i. Larger Fisher score values designate higher discriminating feature power.

The evaluation is done by training and testing a specific classification model that estimates the relevance of a given subset of characteristics. Although filter methods are fast and scalable, they have the disadvantage of ignoring the interaction with the classifier. Usage of nested cross-validation allows one to avoid overoptimistic results. In nonnested feature selection, features are selected based on entire training set and only after that cross-validation is applied. In nested feature selection the best feature selection is performed for each fold [16]. In the first case, the model sees the test set implicitly, while the second case resembles real-life scenarios more precisely [17].

5. Results

The previously described workflow was applied to three specific scenarios:

1. Analysis of the movements in half-second time intervals before the freezing episodes.
2. Analysis of movements in half-second time intervals after the freezing episodes.
3. Analysis of movements in one second around(half second before and after) the freezing episodes.

Each classifier was used with the best 2, 3, 4 and 5 features. Due to the size of the data set, the usage of a larger number of features did not make sense.

5.1 Movements before freezing episode results

Fisher's score feature selection for the movements occurring before the freezing episodes has showed that highest discriminating power have following features: velocity mean, its vertical projection and standard deviation, followed by acceleration mean and its standard deviation.

Feature	Fisher's score
mean of velocity	0,420
mean of velocity vertical projection	0,402
standard deviation of velocity	0,374
mean of acceleration	0,374
standard deviation of acceleration	0.372

Table 1. Best features Fisher's score for movements before the freezing episodes

The number of features have not affected goodness of the models. The accuracy have varied between 0.73 and 0.78, precision 0.76 and 0.78, recall 0.81 and 0.93. SVM and LR have demonstrated best result, but it was just slightly better compared to other classifiers.

	Recall	**F1**	**Accuracy**	**Precision**	**Roc**	**Specificity**
AB	0.89	0.83	0.77	0.78	0.83	0.58
DT	0.85	0.81	0.75	0.78	0.79	0.59
KNN	0.83	0.81	0.75	0.80	0.79	0.59
LR	0.90	0.84	0.78	0.79	0.82	0.54
RF	0.84	0.82	0.76	0.80	0.82	0.60
SVM	0.92	0.84	0.77	0.77	0.78	0.50

Table 2. Classifier results with 5 features for movements before the freezing episodes

5.2 Movements after freezing episode results

Similarly to movements before freezing, the mean and standard deviation of the acceleration with the velocity mean were present in the list of most discriminating features for movements after freezing. Also, their Fisher's score values were around ten per cent higher. The mean of a standard deviation of the vertical acceleration projection filled the remaining spots of the list.

Feature	**Fisher's score**
mean of acceleration	0,570
standard deviation of acceleration	0,555
mean of acceleration vertical projection	0,538
mean of velocity	0,531
standard deviation of acceleration vertical projection	0.531

Table 3. Fisher's score for movements after the freezing episodes

This difference is reflected by the goodness of the classifiers. The accuracy has varied between 0.78 and 0.81, precision 0.82 and 0.85, recall 0.85 and 0.91. LR and SVM classifiers again showed the best results.

	Recall	**F1**	**Accuracy**	**Precision**	**Roc**	**Specificity**
AB	0.89	0.86	0.81	0.83	0.85	0.67
DT	0.90	0.86	0.82	0.82	0.83	0.64
KNN	0.87	0.86	0.82	0.85	0.84	0.71
LR	0.90	0.86	0.81	0.82	0.86	0.64
RF	0.87	0.85	0.80	0.83	0.83	0.68
SVM	0.90	0.86	0.81	0.83	0.84	0.64

Table 4. Classifier results with 5 features for movements after the freezing episodes

5.3 Movements around freezing episode results

For the analysis of the movements around freezing, Fisher's score has shown that mean and standard deviation of acceleration and its vertical projection together with the jerk standard deviation have the highest discriminating power.

Feature	Fisher's score
mean of acceleration	0,775
standard deviation of acceleration	0,763
mean of acceleration vertical projection	0,749
standard deviation of jerk	0,743
standard deviation of vertical acceleration projection	0.741

Table 5. Fisher's score for movements around freezing episodes

Model goodness metrics have varied as follows: the accuracy has varied between 0.74 and 0.8, precision 0.79 and 0.84, recall 0.84 and 0.91. All classifiers showed similar results.

	Recall	F1	Accuracy	Precision	Roc	Specificity
AB	0.90	0.85	0.80	0.82	0.85	0.62
DT	0.90	0.82	0.80	0.80	0.82	0.63
KNN	0.84	0.82	0.74	0.83	0.81	0.74
LR	0.90	0.85	0.79	0.80	0.83	0.61
RF	0.87	0.84	0.80	0.84	0.85	0.65
SVM	0.91	0.85	0.79	0.79	0.80	0.56

Table 6. Classifier results with 5 features for movements around freezing episodes

5.4 Results illustration

For results illustration was made graph of the velocity in one second interval around the freezing episode.

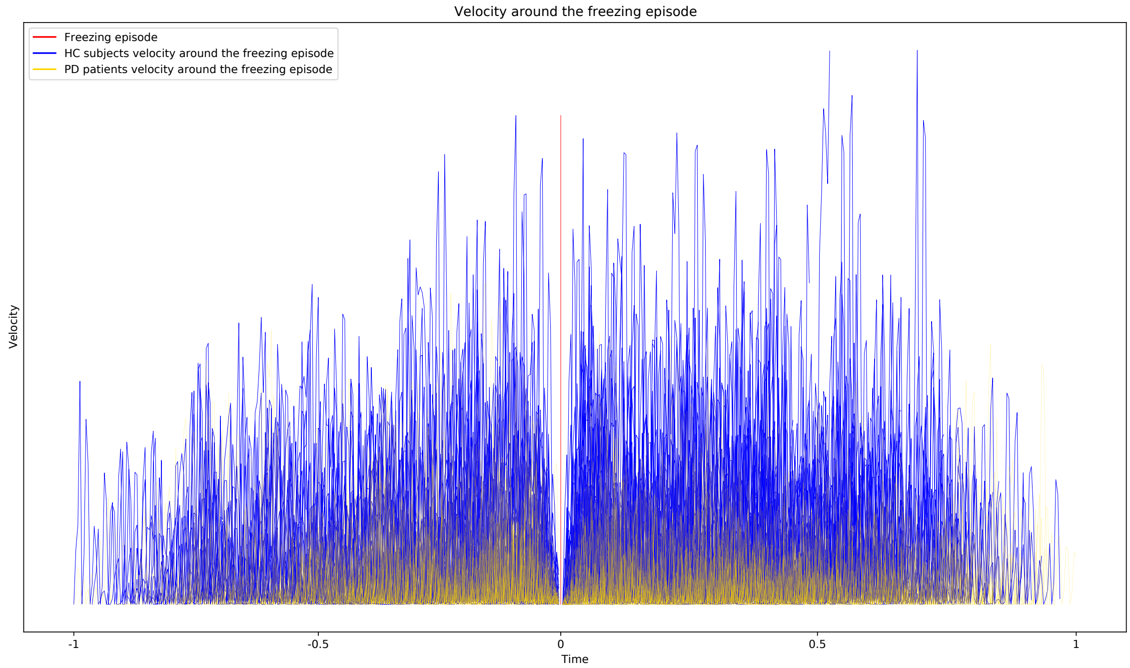

Figure 4. Velocity around freezing episodes.

In Figure 4 blue lines represent the velocities of HC subjects and yellow lines represent velocities of PD patients observed around freezing episodes in the motions. The red line in the middle shows the freezing point. It is noticeable from the figure that mean and standard deviation values are distinguishable between HC and PD groups.

6. Illiteracy analysis

To analyse re-usability possibilities of the developed during research approach, it was used with the new DraWriteUniPampa dataset. The new data set consists of 3 drawing and 1 writing tests, which were made by 71 volunteers with different levels of education. The tests of each type were divided into 2 groups. Into first group belonged 52 illiterate test subjects and second group consisted of 19 subjects with proper writing knowledge.

The data collection process was similar to the acquisition of DraWritePD datasets, which was described in Chapter 3 of the thesis. Only differences are that for volunteers was used another group of people without PD, who live in Brazil, and instead of iPad was used Microsoft Surface tablet. The data processing and analysis process was the same and is described in Chapter 4.

Analysis was performed only for movements in one second around (half a second before and after) the freezing episodes. Other scenarios in Chapter 5 were skipped because there were no significant differences between the results of the sentence writing test.

The data acquisition, processing and analysis were performed with the permission of the ethics committee and according to the data handling and privacy protection laws of the countries involved. For data acquisition, the permit was issued by Universidade da regiao de Joinville (UNIVILLE) nr. 4.787.687 on June 17, 2021. For data handling and processing by Tallinn University 6-5.1/14 on May 12, 2021. Participation in the experiment was voluntary.

6.1 Name writing test

Writing test possibilities were strongly limited due to fact, that part of volunteers are illiterate and do not have writing knowledge. In only possible to conduct writing test, volunteers were asked to write their own name on the tablet screen. Figure 5 shows one completed test example.

Figure 5. Name written on the screen of a tablet.

Maximum alpha angle value became the most discriminating feature in the name writing test with big gap in Fisher's score from other features. The masses of the derivatives of position took the next four places with small differences in score between each other.

Feature	**Fisher's score**
maximum alpha angle	0.822
velocity mass	0.168
acceleration mass	0.161
jerk mass	0.159
snap mass	0.158

Table 7. Fisher's score for movements during name writing test

The accuracy has varied between 0.70 and 0.74, precision 0.74 and 0.78, recall 0.86 and 0.99, but specificity values were low and varied between 0.13 and 0.6 based on model.

	Recall	**F1**	**Accuracy**	**Precision**	**Roc**	**Specificity**
AB	0.98	0.84	0.74	0.75	0.73	0.24
DT	0.99	0.84	0.74	0.75	0.73	0.18
KNN	0.88	0.81	0.71	0.78	0.66	0.58
LR	0.94	0.84	0.74	0.76	0.73	0.22
RF	0.86	0.80	0.70	0.76	0.68	0.41
SVM	0.98	0.85	0.75	0.74	0.57	0.14

Table 8. Classifier results with 5 features for movements around freezing during name writing test

6.2 Bowl drawing test

Volunteers were asked to draw figure similar to the bowl. Example of the test is present in Figure 6

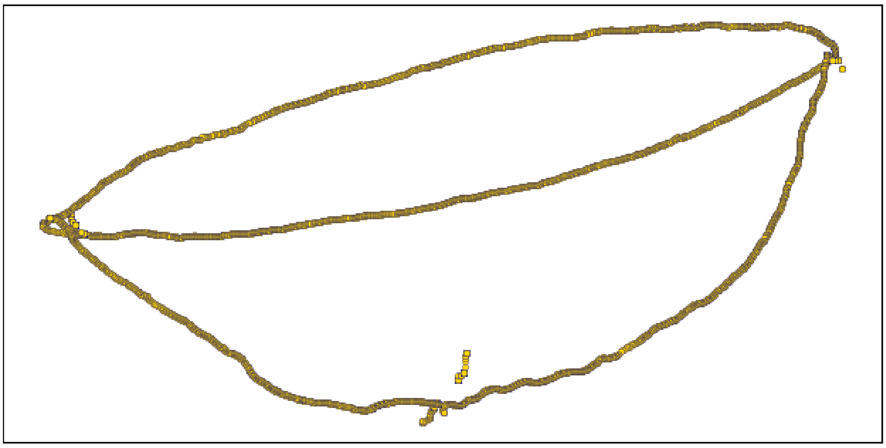

Figure 6. Bowl on the screen of a tablet.

Maximum alpha angle value also took lead in bowl test, but Fisher's score value was more than 2 times smaller compared to name test. Other best features were connected to altitude and had 6 times smaller values compared to most discriminating feature.

Feature	**Fisher's score**
maximum alpha angle	0.328
standard deviation of altitude angle change	0.059
maximum altitude angle	0.052
horizontal trajectory	0.049
altitude angle change mass	0.047

Table 9. Fisher's score for movements during bowl drawing test

Recall values increased and lay between 0.96 and 0.99, but specificity dropped to 0.10 and

0.20. F1, accuracy and precision values increased by 0.05 to 0.1 on average compared to name test results.

	Recall	**F1**	**Accuracy**	**Precision**	**Roc**	**Specificity**
AB	0.99	0.92	0.86	0.86	0.74	0.12
DT	0.98	0.91	0.85	0.86	0.67	0.13
KNN	0.96	0.90	0.83	0.85	0.66	0.12
LR	0.98	0.91	0.85	0.85	0.73	0.10
RF	0.96	0.91	0.85	0.87	0.73	0.20
SVM	0.99	0.92	0.85	0.85	0.62	0.05

Table 10. Classifier results with 5 features for movements around freezing during bowl drawing test

6.3 Multicube drawing test

Test subjects were asked to draw parallelepiped figure. Example is shown in Figure 7.

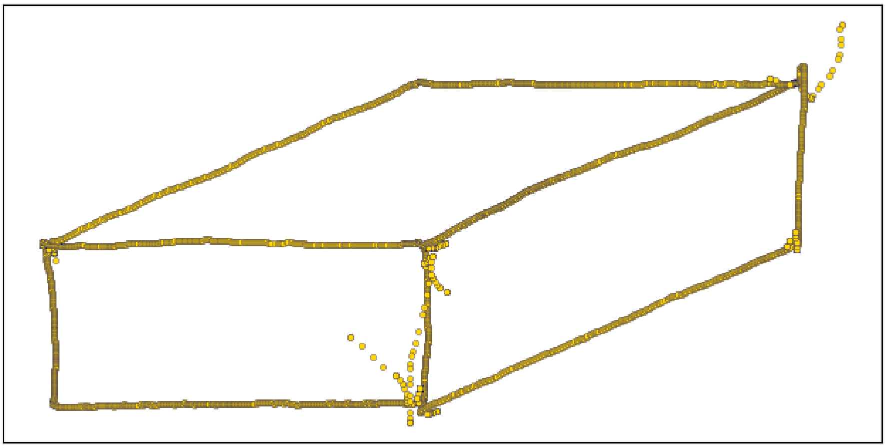

Figure 7. Multicube on the screen of a tablet.

Maximum alpha angle value had biggest Fisher's score value for multicube test too. Similar to name test into other best features belonged derivatives of position, but values of all features were 4-5 times less, compared to name test.

Feature	Fisher's score
maximum alpha angle	0.203
jerk mass	0.033
snap mass	0.033
acceleration mass	0.033
crackle mass	0.033

Table 11. Fisher's score for movements during multicube drawing test

Recall have incraesed for 1 for some models, but specificity in return dropped to almost 0. F1, accuracy and precision had similar results compared with bowl test.

	Recall	F1	Accuracy	Precision	Roc	Specificity
AB	1.00	0.91	0.84	0.84	0.64	0.00
DT	1.00	0.91	0.84	0.84	0.62	0.00
KNN	0.96	0.90	0.83	0.85	0.59	0.07
LR	0.99	0.91	0.84	0.84	0.63	0.08
RF	0.95	0.89	0.81	0.85	0.58	0.07
SVM	1.00	0.91	0.84	0.84	0.49	0.00

Table 12. Classifier results with 5 features for movements around freezing during multicube drawing test

6.4 Hexagon drawing test

During hexagon drawing test, subjects were drawing figure, what resembles honeycombs. In Figure 8 is captured one of the acquired examples.

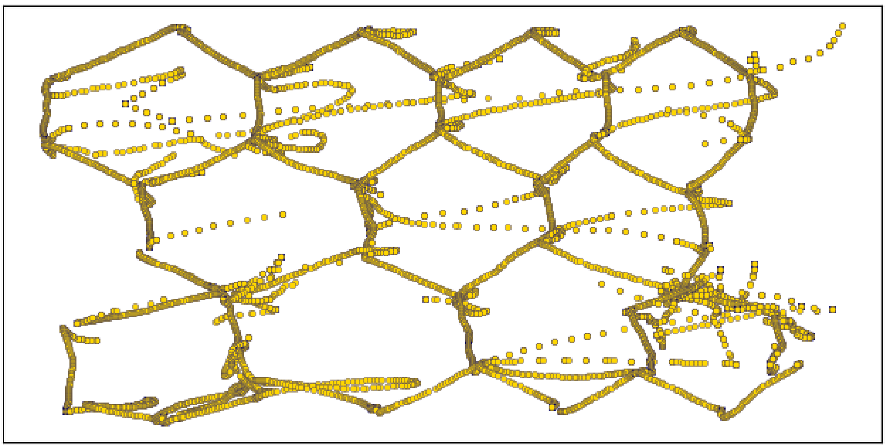

Figure 8. Hexagon on the screen of a tablet.

All best features by Fisher's score are connected to azimuth and altitude parameters. For all previously analysed tests best feature had much bigger value, compared for features on

2nd-5th places, but for hexagon drawing test all features have low discrimination power with Fisher's score value less than 0.1.

Feature	**Fisher's score**
azimuth angle velocity mass	0.071
azimuth angle change mass	0.068
maximum azimuth angle velocity	0.060
azimuth angle acceleration mass	0.059
altitude angle velocity mass	0.057

Table 13. Fisher's score for movements during hexagon drawing test

However low Fisher's score values did not affect model goodness results and values are similar to previous test cases. Recall values vary between 0.88 and 0.97, accuracy 0.69 and 0.75, precision 0.73 and 0.77. Specificity values became better compared to multicube results, but still are quite low and lay between 0.13 and 0.31.

	Recall	**F1**	**Accuracy**	**Precision**	**Roc**	**Specificity**
AB	0.96	0.84	0.74	0.75	0.63	0.20
DT	0.94	0.83	0.74	0.75	0.62	0.23
KNN	0.88	0.80	0.69	0.74	0.57	0.21
LR	0.97	0.84	0.74	0.75	0.61	0.18
RF	0.93	0.84	0.75	0.77	0.71	0.31
SVM	0.97	0.84	0.73	0.73	0.57	0.13

Table 14. Classifier results with 5 features for movements around freezing during hexagon drawing test

7. Discussion

The analysis of the sentence writing test with patients with PD was already done in the research [8]. Subjects were asked to write Chezch *Tramvaj dnes už nepojede* which means *The tram will not go today*. It used similar techniques for sentence analysis, but examined in test sentence was written in another language, had different length, and was combined with other drawing tests. Therefore, comparison of results with current work is problematic and could give false conclusions. If not to take this issue into account, the goodness of the model for movements before freezing is similar to the results in [8]. The goodness of ML models for movements after the freezing episode even exceeds previously acquired results.

The sentences from the same data set, which is used in the present thesis, were analysed in the [11] work. The main goal of the research was to split the test not by freezing, but by letters and to analyse them individually. Usage of the same data set allows for more precise results comparison. The goodness metrics of the model in [11] lie in the same range and have the following results. The precision ranged between 0.73 and 0.82, the precision between 0.71 and 0.91, and the recall between 0.61 and 0.93 (excluding the SVM and KNN models, which showed poor performance). The most frequent features with the highest Fisher's score were the mean velocity and the mean acceleration mass of the angular change.

The results of the analysis of freezing episodes do not surpass micrografia or the use of other tests for the diagnosis of PD, but the developed approach could be used as valuable additional support for disease diagnostics. Research findings show that usage of small part of the test around freezing episode can be as informative as usage of the whole test. This consequence was also present in [21].

One more important research finding is that the freezing analysis approach could be used for other purposes. Currently, it was tested only to confirm differences in performance during name writing and figures drawing tests of illiterate and educated volunteers. It is impossible to say how generic and for which tasks the developed approach is suitable, but acquired results indicate good potential for re-usability possibilities.

8. Conclusion

The attention of the present research has been focused on the stylus tip movements observed half second before, half second after and second around the hand freezing episodes. The application was developed to find freezing episodes, extract motion data around the stop, find best features and use them with most common ML models. In addition developed approach has shown re-usability possibilities for other goals. It was used on different dataset to confirm difference in drawing and writing tests between educated and illiterate subjects.

It was demonstrated that post-freezing movements may provide enough information for the entire test. The quality metrics of the trained ML classifiers are comparable and sometimes even better than those reported in the literature.

The results of the thesis indicated that the main goals were successfully achieved, but there are still many possible improvements. Present research can evolve in different ways:

- Better freezing recognition algorithm could be developed to exclude actual freezing what was caused by PD from intentional stops
- Currently developed approach could be used on other datasets to verify future possibilities of method re-usability.
- More attention could be paid to the time frames of the freezing neighbourhood. How an increase/decrease of the time interval around the stop in different tests would affect ML classifiers.

Acknowledgements

This thesis is partially support by the Estonian Research Council ETAG grand PRG 2100.

References

[1] Lorraine V Kalia and Anthony E Lang. "Parkinson's disease". In: *The Lancet* 386.9996 (2015), pp. 896–912. ISSN: 0140-6736.

[2] Elan D Louis and Duarte G Machado. "Tremor-related quality of life: a comparison of essential tremor vs. Parkinson's disease patients". In: *Parkinsonism & related disorders* 21.7 (2015), pp. 729–735.

[3] Wu Wang et al. "Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning". In: *IEEE Access* 8 (2020), pp. 147635–147646. DOI: 10.1109/ACCESS.2020.3016062.

[4] Esther Smits et al. "Standardized Handwriting to Assess Bradykinesia, Micrographia and Tremor in Parkinson's Disease". In: *PloS one* 9 (May 2014). DOI: 10.1371/journal.pone.0097614.

[5] Ahmed A. Moustafa et al. "Motor symptoms in Parkinson's disease: A unified framework". In: *Neuroscience and Biobehavioral Reviews* 68 (2016), pp. 727–740. ISSN: 0149-7634. DOI: https://doi.org/10.1016/j.neubiorev.2016.07.010.

[6] Evelien Nackaerts et al. "Validity and reliability of a new tool to evaluate handwriting difficulties in Parkinson's disease". In: *PloS one* 12.3 (2017), e0173157.

[7] C. Marquardt and N. Mai. "A computational procedure for movement analysis in handwriting". In: *Journal of Neuroscience Methods* 52.1 (1994), pp. 39–45. ISSN: 0165-0270. DOI: http://dx.doi.org/10.1016/0165-0270(94)90053-1.

[8] Peter Drotár et al. "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease". In: *Artificial Intelligence in Medicine* 67 (2016), pp. 39–46. ISSN: 0933-3657. DOI: https://doi.org/10.1016/j.artmed.2016.01.004.

[9] S. Nõmm et al. "Detailed Analysis of the Luria's Alternating SeriesTests for Parkinson's Disease Diagnostics". In: *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2018, pp. 1347–1352. DOI: 10.1109/ICMLA.2018.00219.

[10] Elli Valla et al. "Tremor-related feature engineering for machine learning based Parkinson's disease diagnostics". In: *Biomedical Signal Processing and Control* 75 (2022), p. 103551. ISSN: 1746-8094. DOI: https://doi.org/10.1016/j.bspc.2022.103551.

[11] Aleksei Netšunajev et al. "Parkinson's Disease Diagnostics Based on the Analysis of Digital Sentence Writing Test". In: *Vietnam Journal of Computer Science* 8.04 (2021), pp. 493–512.

[12] Elke Heremans et al. "Amplitude Manipulation Evokes Upper Limb Freezing during Handwriting in Patients with Parkinson's Disease with Freezing of Gait". In: *PLOS ONE* 10 (Nov. 2015), pp. 1–13. DOI: 10.1371/journal.pone.0142874. URL: https://doi.org/10.1371/journal.pone.0142874.

[13] A Wagle Shukla et al. "Micrographia and related deficits in Parkinson's disease: a cross-sectional study". In: *BMJ open* 2.3 (2012). DOI: https://doi.org/10.1136/bmjopen-2011-000628.

[14] AWA Van Gemmert, Charles Howard Adler, and GE Stelmach. "Parkinson's disease patients undershoot target size in handwriting and similar tasks". In: *Journal of Neurology, Neurosurgery & Psychiatry* 74.11 (2003), pp. 1502–1508.

[15] Vassili Gorbatsov et al. "Machine Learning Based Analysis of the Upper Limb Freezing During Handwriting in Parkinson's Disease Patients". In: *IFAC-PapersOnLine* 55.29 (2022), pp. 91–95.

[16] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning. 2nd Edition*. Springer Series in Statistics. Springer, 2002.

[17] Maria Teresa Angelillo et al. "Performance-driven handwriting task selection for Parkinson's disease classification". In: *International Conference of the Italian Association for Artificial Intelligence*. Springer. 2019, pp. 281–293.

[18] S. Nõmm et al. "Recognition and Analysis of the Contours Drawn during the Poppelreuter's Test". In: *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2016, pp. 170–175. DOI: 10.1109/ICMLA.2016.0036.

[19] Santiago Perez-Lloret et al. "Prevalence, determinants, and effect on quality of life of freezing of gait in Parkinson disease". In: *JAMA neurology* 71.7 (2014), pp. 884–890.

[20] C. Aggarwal. *Data Mining*. Springer, 2015.

[21] Sven Nomm et al. "Determining Necessary Length of the Alternating Series Test for Parkinson's Disease Modelling". In: *2019 International Conference on Cyberworlds (CW)*. IEEE. 2019, pp. 261–266.

Appendix 1 – Non-Exclusive License for Reproduction and Publication of a Graduation Thesis[1]

I Vassili Gorbatšov

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis “Motion Freezing Analysis in Drawing and Writing Tests ”, supervised by Sven Nõmm
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

27.05.2024

[1]The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.