# Automatic Assessment of Bradykinesia in Parkinson's Disease Using Tapping Videos

Matjaž Zupanič
matjaz.zupanic1@gmail.com
Faculty of Computer and Information Science,
University of Ljubljana
Večna pot 113
SI-1000 Ljubljana, Slovenia

Dejan Georgiev
dejan.georgiev@kclj.si
Faculty of Medicine,
University of Ljubljana
Vrazov trg 2
SI-1000 Ljubljana, Slovenia

Jure Žabkar
jure.zabkar@fri.uni-lj.si
Faculty of Computer and Information Science,
University of Ljubljana
Večna pot 113
SI-1000 Ljubljana, Slovenia

## ABSTRACT

Parkinson's disease is a chronic neurodegenerative illness that severely affects the everyday life of a patient. The severity of Parkinson's disease is assessed using the MDS-UPDRS scale. In this study, we explore the feasibility of automatically evaluating bradykinesia, a key symptom of Parkinson's disease, from tapping videos recorded on smartphones in everyday settings. We collected a dataset of 183 tapping videos, from 91 individuals. Videos were assessed by neurologist into 5 classes of the MDS-UPDRS scale. For data extraction, we employed MediaPipe Hand, which provides a time series of hand skeleton movements. The data was preprocessed to eliminate noise and subsequently used for either feature construction or directly in neural networks. Utilizing manually created features in a multilayer perceptron classifier resulted in 61 % accuracy and an F1 score of 0.61 on our test set. Employing a fully convolutional network, we improved the accuracy to 78 % and the F1 score to 0.75. Additionally, we developed the tool for visualising tapping and displaying key data, providing detailed insights into tapping patterns.

## KEYWORDS

Parkinson's disease, Finger tapping test, Machine learning

## 1 INTRODUCTION

Parkinson's disease (PD) is a chronic neurodegenerative condition that profoundly impacts daily life. PD affects 1-2 % [23] of the population over the age of 65. Currently, there are more than 1.2 million cases in Europe [4] and this number is forecast to double in the near future due to the demographic problem of an aging population. Its etiology remains incompletely understood, yet researchers suggest that a combination of genetic and environmental factors contributes to its development. Factors such as exposure to polluted air, pesticides, heavy metals, and head injuries have been associated with an increased risk of Parkinson's disease. The most common symptoms include bradykinesia, which is also the main symptom, tremor, rigidity, impaired postural reflexes, and dementia. There are also numerous other symptoms that can accompany the disease, such as sleep disturbances, depression, loss of smell, and fatigue.

The standardized MDS-UPDRS [6] scale is used to assess the stage of Parkinson's disease. It consists of 4 sections that evaluate both motor and non-motor issues experienced by patients. The finger-tapping test is used to evaluate the severity of bradykinesia. This test involves asking the individual to tap their index finger and thumb as quickly as possible with a maximal span, assessing the number of pauses, time taken, decrease in amplitude, and slowing of speed, all contributing to the final score. It was estimated that up to 25 % of clinical diagnoses of PD are incorrect, due to lack of experience or attention during tests [3].

## 2 RELATED WORK

First automated systems for PD detection were based on wearable sensors like gyros and accelerometers [5, 17, 20, 22] or on electromyography sensors [11, 28]. The main issues with sensors are that they are commercially unavailable, require precise placement, and can interfere with tapping test. Therefore, some researchers have utilized keyboard tapping [1, 19] or tapping on a smartphone screen [7, 8] for data acquisition. Sadikov et al. [21] collected data using digital spirometry, where participants traced an Archimedes spiral on a touchscreen device. Advances in hardware and software have made computer vision combined with machine learning a viable alternative for PD recognition, allowing for home testing using a computer or smartphone. Lainscsek et al. [13] used a nonlinear delay differential equation, with the structure selected by a genetic algorithm. While other researchers used machine learning techniques, most focused on manual feature creation and utilized these features in classification models [10, 18, 29, 30] like support vector machines (SVM) and random forests (RF), or regression models [9] like support vector regression (SVR) and XGBoost. Others employed neural networks (NN) to automatically learn patterns from time series data [2, 14]. Researchers used segmentation neural networks or optical flow for hand data extraction, with MediaPipe Hand [16] becoming popular in later works. Due to limited data, most studies combined some classes, and only a few performed full-scale classification [2, 9, 14, 30].

## 3 METHODOLOGY

First, we collected and labeled data and preprocessed it to eliminate noise. We explored various approaches. Initially, we manually extracted features from the time series data of the hand skeleton and applied SVM, multi-layer perceptron (MLP), and random forest (RF). Later, we employed a fully convolutional neural network (FCN), illustrated in Figure 2.

### 3.1 Dataset

Since available datasets of tapping videos were not publicly accessible, we assembled our own database. From each participant, we collected two videos: one of the left-hand tapping and another of the right-hand, as they are independent from each other and have their own MDS-UPDRS score. Videos were recorded at a resolution of 3840 x 2160 or 1920 x 1080 at 60 or 30 fps. All PD patients were recorded in a clinical setting, while some participants of the healthy group were recorded in various other environments. We removed videos from the database that had significant tilts, where the hand was not visible throughout the entire action, and those of participants who scored higher than 0 by MDS-UPDRS but did not have confirmed PD. In total, we compiled 183 videos from 91 different individuals. The distribution of data between classes can be observed in Table 1.

| MDS-UPDRS score | 0 | 1 | 2 | 3 | 4 | Sum |
|---|---|---|---|---|---|---|
| Number of videos | 49 | 51 | 53 | 23 | 7 | 183 |
| Percentage % | 26 | 28 | 29 | 13 | 4 | 100 |

**Table 1: Number of videos in each class.**

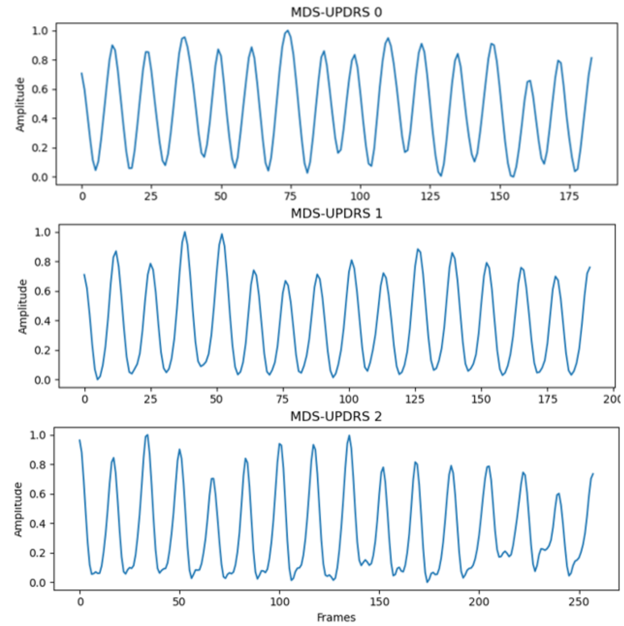


**Figure 1: Amplitude graphs of finger tapping.**

## 3.2 Preprocessing

Since datasets were collected without any professional equipment we were dealing with different illuminations, angles, camera tilts, distances between camera and hand, noise, and motion blur. Videos were cut, so only hand is visible, due to privacy and faster processing. For hand skeleton extraction, we used Mediapipe Hand. We then computed Euclidean distances between the endpoints of the thumb and index finger to obtain time series data. Occasionally, there were single-frame misses where Mediapipe detected previous and next frames but missed the current one. In such cases, we applied linear interpolation to fill in the missing value. However, we avoided interpolating longer gaps to preserve tapping integrity. To reduce noise caused by inaccurate detections from MediaPipe, we implemented a combination of low-pass and moving average filter. Filtering also helped to eliminate tremor which is not part of MDS-UPDRS, but masked tapping. We balanced filtering and signal preservation using weaker filters. The Low-pass filter addressed high-frequency noise from tremors and MediaPipe Hand's misalignment, which caused slight shifts between frames even when the hand was still. We implemented the Butterworth low-pass filter due to its flat response. The cutoff frequency was set uniquely for each input based on a specified percentage of influence (k) (equation 1). Additionally, a moving average with a window size of 5 was used to further smooth the data and reduce erroneous detections, such as spikes.

$$f_{cutoff} = f_{max} + \frac{\text{band width} * k}{2} \tag{1}$$

We restricted tapping sequences to a maximum of 15 taps, as the MDS-UPDRS scale does not require longer sequences and participants may tire during extended sessions. We also applied min-max normalization to account for varying distances between the camera and the hand. The finalised graphs of the processed data are depicted in Figure 1.

## 3.3 Feature Engineering

In the **1st method** we created a larger number of features following the MDS-UPDRS scale to comprehensively describe finger tapping. In addition to Euclidean distances between the endpoints of the thumb and index finger, we included distances between the last joints and absolute wrist movement as additional time series data. We hypothesized that these metrics could provide supplementary but limited insights into finger tapping, although movements in these areas are typically less pronounced. The time series of distances represents the amplitude spectrum, from which we derived 4 additional spectra: velocity, acceleration, frequency, and spectrum of amplitude peaks. Additionally, we included the spectrum of absolute wrist movement. From these 6 spectra, we extracted 193 statistical features. Our goal was to capture dependencies at global and local levels, describing hesitations, slowdowns, amplitude decreases, data distributions, tapping energy, and other characteristics.

In **2nd method** we designed features closely aligned with the MDS-UPDRS scale, categorizing them into 3 parts reflecting its criteria. The first part assesses hesitations and freezes, the second part measures reduced speed and the third part evaluates decreased amplitudes. In our final analysis, we utilized 145 features, with 90 being coefficients from linear regressions to assess reductions in amplitudes and velocities of finger tapping. These coefficients were derived from local maxima of amplitudes and velocities. The remaining 55 features were derived from amplitudes, velocities, their extremes, and autocorrelated velocities and amplitudes.

## 3.4 Neural network

We tested the FCN presented by Li et al. [14], using preprocessed time series data directly as input. Since the selected FCN is limited to processing equally long inputs, we padded our time series data with 0 at the end. We later modified the FCN by adding convolutional layers, dropout layers, early stopping, and adjusted input layers to handle 2D inputs consisting of amplitudes and velocities. The architecture of our extended FCN is shown in Figure 2.

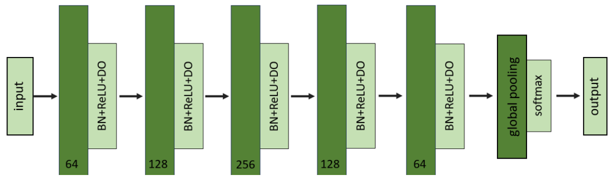


**Figure 2: FCN for classification of bradykinesia.**

## 4 EVALUATION

We created a visualization tool for analyzing finger-tapping dynamics, featuring a built-in video player with a MediaPipe Hand skeleton overlay. Velocity and amplitude graphs on the right side indicate the current frame with a vertical red line (Figure 3). Taps are denoted by red dots and the tool includes a freeze labelling option. This comprehensive and interactive analysis of finger-tapping dynamics is particularly beneficial for neurologists in diagnosing and monitoring motor disorders.

All tests were conducted using 10-fold cross-validation due to the relatively small dataset. In Methods 1 and 2, we used SelectKBest [26] for feature selection. As score functions we tried ANOVA F-test [24] and mutual information for a discrete target [25], with the latter performing better overall. By experimenting with various feature counts, we identified the optimal number that maximized the model's F1 score. Results are detailed in Table 2.

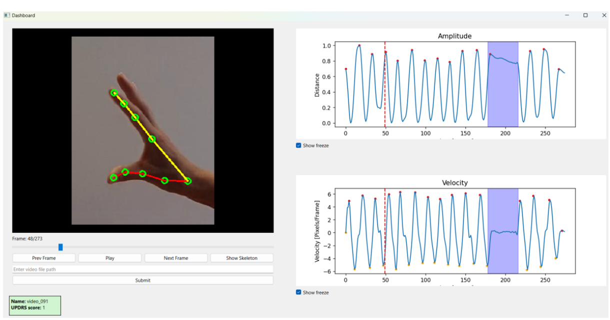


**Figure 3: Tool for detailed visualization of finger tapping and display of key data.**

| Model | Accuracy % | F1 % | Precision % | Recall % |
|---|---|---|---|---|
| FCN | 72 | 62 | 84 | 63 |
| **FCN +** | **77** | **75** | **88** | **75** |
| Method 1 | 61 | 62 | 67 | 58 |
| Method 2 | 60 | 55 | 67 | 58 |

**Table 2: Results were obtained via 10-fold cross-validation. FCN refers to Li et al.'s neural network [14], while FCN+ denotes our modified version (Figure 2). MLP achieved the best result for Method 1 and RF for Method 2.**

## 5 CONCLUSION

Our data set was diverse, assembled by tapping videos of different people among all MDS-UPDRS classes. Since the dataset was collected without any professional equipment we were dealing with different illuminations, angles, camera tilts, distances between hand and camera, noise, and motion blur. We opted for milder filtering to strike a balance between noise reduction and preserving signal integrity. Our class with MDS-UPDRS score of 4 had a limited population. Therefore, to obtain more precise results, it would be essential to expand the dataset. Additionally, more data could be utilized for an unseen test dataset. Due to the low capture speed and fast movement of fingers, motion blur was present in the videos. To address this issue, we employed NNs for motion blur removal. We tested two different NNs: Ghost-DeblurGAN [15] and PDV_net [27]. However, both methods introduced artifacts in the frames, prompting us to discontinue their use. We also experimented with upscaling the resolution to 200 % of the original size using Video2x [12]. This aimed to enhance image clarity, potentially improving MediaPipe Hand's skeleton detection precision and reducing noise. Testing on a smaller upscaled subset showed minimal differences in classification performance but significantly increased processing time, prompting us to abandon this approach due to time constraints.

Similarities were observed across different classes of tapping, as shown in Figure 1, where the graphs appear similar. Various factors contribute to the overlapping of classes. Small differences in tapping styles between adjacent classes mean even minor decreases in velocity or amplitude can significantly impact the final classification. Approximately 83 % and 73 % of misclassifications using methods 1 and 2 differed from the reference tapping scores by exactly 1 class. Normalization contributes to reduced differentiation between classes by masking tapping instances with low amplitudes. However, it is necessary to account for variations in recording distances and resolutions. Additionally, recording angles can distort actual finger distances, leading to misleading data representation. Another factor for class overlap is the possibility of errors in video assessments. Hence, we plan to involve another neurologist to cross-validate our dataset to eliminate potential human errors. However, within the same class, tapping behaviors can vary significantly. For example, in class 4, participants often showed varying abilities: some managed to perform a few taps despite severe difficulties, while others were unable to tap at all. Overall SVM performed the worst in methods 1 and 2, while RF and MLP produced the best results, likely because of their superior ability to capture and model complex patterns in the data. Therefore, Method 1 performed better overall metrics than Method 2, which also had 10 % more misclassifications that differed by more than 1 class from the reference value. Using FCN from Figure 2 led to significantly better performance as shown in Table 2. With manual time-invariant feature creation, it is challenging to capture all unique patterns at various scales in a tapping sequence of around 400 frames. FCN excels at extracting both local and global dependencies from time series using convolutional filters of various sizes, capturing and interpreting information at multiple scales. These features make the FCN superior in analyzing tapping, despite requiring longer training times, larger databases, and more powerful systems. When comparing our study with related research, direct comparisons may be challenging due to the use of different datasets. When comparing our Methods 1 and 2 using the method by Yu et al. [30], who derived features based on MDS-UPDRS, we achieved lower scores. They reported 80 % accuracy and 79 % recall on a test set of only 15 videos recorded as close to a 90-degree angle as possible. Frame interpolation they used might distort tapping details with artificial data, risking the reliability of their classification outcomes. The FCN presented by Li and colleagues [14] achieved 72 % accuracy on our dataset, compared to the 80 % reported by the authors. We attribute the slightly lower classification performance on our data to its complexity and heterogeneity, over 37 % smaller size, and the use of 10-fold cross-validation compared to their 5-fold approach. However, by enhancing the FCN (Figure 2) we improved prediction accuracy to 77 %. For future work, we plan to explore graph neural networks (GNN) for their capability to handle all data points extracted by MediaPipe Hand. Alam et al. [2] reported 81 % accuracy and F1 score of 0.81 on their test set using a GNN. We also plan to explore regression models that can predict continuous severity scores, offering a more detailed evaluation of bradykinesia. Islam et al. [9] investigated SVR, LightGBM, and XGBoost, achieving up to 55 % accuracy. This is lower than the 61 % accuracy we achieved with Method 1, possibly due to their larger database of 489 videos, less effective preprocessing and a feature set of 65 features that may not fully capture tapping dynamics.

## REFERENCES


[1] Warwick R Adams. 2017. High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing. *PLOS ONE* 12, 11 (11 2017), 1–20. https://doi.org/10.1371/journal.pone.0188226

[2] Zarif U Alam, Saiful Islam, Ehsan Hoque, and Saifur Rahman. 2023. PULSAR: Graph based Positive Unlabeled Learning with Multi Stream Adaptive Convolutions for Parkinson's Disease Recognition. https://doi.org/10.48550/ARXIV.2312.05780

[3] Nin P S Bajaj, Vamsi Gontu, James Birchall, James Patterson, Donald G Grosset, and Andrew J Lees. 2010. Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study. *Journal of Neurology, Neurosurgery & Psychiatry* 81, 11 (2010), 1223–1228. https://doi.org/10.1136/jnnp.2009.193391

[4] Parkinson's Europe. 2024. Parkinson's Statistics. https://parkinsonseurope.org/facts-and-figures/statistics/ Accessed: 7-3-2024.

[5] Joseph P Giuffrida, David E Riley, Brian N Maddux, , and Dustin A Heldman. 2009. Clinically deployable Kinesi technology for automated tremor assessment. *Movement Disorders* 24, 5 (2009), 723–730. https://doi.org/10.1002/mds.22445

[6] Christopher G Goetz, Barbara C Tilley, Stephanie R Shaftman, Glenn T Stebbins, Stanley Fahn, Pablo Martinez-Martin, Werner Poewe, Cristina Sampaio, Matthew B Stern, Richard Dodel, Bruno Dubois, Robert Holloway, Joseph Jankovic, Jaime Kulisevsky, Anthony E Lang, Andrew Lees, Sue Leurgans, Peter A LeWitt, David Nyenhuis, Warren C Olanow, Olivier Rascol, Anette Schrag, Jeanne A Teresi, Jacobus J van Hilten, and Nancy LaPelle. 2008. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Movement disorders: official journal of the Movement Disorder Society* 23, 15 (2008), 2129–2170.

[7] Dimitrios Iakovakis, Stelios Hadjidimitrio, Vasileios Charisis, Sevasti Bostantjopoulou, Zoe Katsarou, Lisa Klingelhoefer, Heinz Reichmann, Sofia B Dias, José A Diniz, Dhaval Trivedi, Ray K Chaudhuri, and Leontios J Hadjileontiadis. 2018. Motor impairment estimates via touchscreen typing dynamics toward Parkinson's disease detection from data harvested in-the-wild. *Frontiers ICT* 5 (2018).

[8] Dimitrios Iakovakis, Stelios Hadjidimitriou, Vasileios Charisis, Sevasti Bostantzopoulou, Zoe Katsarou, and Leontios J Hadjileontiadis. 2018. Touchscreen typing-pattern analysis for detecting fine motor skills decline in early-stage Parkinson's disease. *Scientific Reports* 8, 1 (2018).

[9] Saiful Islam, Wasifur Rahman, Abdelrahman Abdelkader, Sangwu Lee, Phillip T Yang, Jennifer L Purks, Jamie L Adams, Ruth B Schneider, Earl R Dorsey, and Ehsan Hoque. 2023. Using AI to measure Parkinson's disease severity at home. *npj Digital Medicine* 6, 156 (2023). https://doi.org/10.1038/s41746-023-00905-9

[10] Jacek Jakubowski, Anna P Chromik, Jolanta Chmielinska, Monika Nojszewska, and Anna K Pruszczyk. 2023. Application of imaging techniques to objectify the Finger Tapping test used in the diagnosis of Parkinson's disease. *Bulletin of the Polish Academy of Sciences. Technical Sciences* 71 (2023), art. no. e144886. https://doi.org/10.24425/bpasts.2023.144886

[11] Hyoseon Jeon, Woongwoo Lee, Hyeyoung Park, Hong J Lee, Sang K Kim, Han B Kim, Beomseok Jeon, and Kwang S Park. 2017. Automatic Classification of Tremor Severity in Parkinson's Disease Using a Wearable Device. *Sensors (Basel)* 17, 9 (Sept. 2017).

[12] k4yt3x. 2024. Video 2x. https://github.com/k4yt3x/video2x Dostopano: 16-02-2024.

[13] Claudia Lainscsek, Peter Rowat, Luis Schettino, Dongpyo Lee, D Song, Cristophe Letellier, and Howard Poizner. 2012. Finger tapping movements of Parkinson's disease patients automatically rated using nonlinear delay differential equations. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 22, 1 (2012). https://doi.org/10.1063/1.3683444

[14] Zhu Li, Lu Kang, Miao Cai, Xiaoli Liu, Yanwen Wang, and Jiayu Yang. 2022. An Automatic Evaluation Method for Parkinson's Dyskinesia Using Finger Tapping Video for Small Samples. *Journal of Medical and Biological Engineering* 42, 3 (1 2022), 351–363. https://doi.org/10.1007/s40846-022-00701-y

[15] Yibo Liu, Amaldev Haridevan, Hunter Schofield, and Jinjun Shan. 2022. Application of Ghost-DeblurGAN to Fiducial Marker Detection. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 6827–6832. https://doi.org/10.1109/IROS47612.2022.9981701

[16] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo L Chang, Ming G Yong, Juhyun Lee, Wan T Chang, Wei Hua, Manfred Georg, and Matthias Grundmann. 2019. MediaPipe: A Framework for Building Perception Pipelines. *ArXiv* abs/1906.08172 (2019). https://doi.org/10.48550/arXiv.1906.08172Focustolearnmore

[17] ALexander Meigal, Saara Rissanen, Mika P Tarvainen, Stefanos Georgiadis, Pasi A Karjalainen, Olavi Airaksinen, and Markku Kankaanpää. 2012. Linear and nonlinear tremor acceleration characteristics in patients with Parkinson's disease. *Physiological measurement* 33, 3 (2012), 395.

[18] Adonay S Nunes, Nataliia Kozhemiako, Christopher D Stephen, Jeremy D Schmahmann, Sheraz Khan, and Anoopum S Gupta. 2022. Automatic Classification and Severity Estimation of Ataxia From Finger Tapping Videos. *Frontiers in Neurology* 12 (2022).

[19] Atemangoh B Peachap, Daniel Tchiotsop, Valérie Louis-Dorr, and Didier Wolf. 2020. Detection of early Parkinson's disease with wavelet features using finger typing movements on a keyboard. *SN Applied Sciences* 2, 10 (2020).

[20] Cameron N Riviere, Stephen G Reich, and Nitish V Thakor. 1997. Adaptive Fourier modeling for quantification of tremor. *Journal of Neuroscience Methods* 74, 1 (1997), 77–87. https://doi.org/10.1016/S0165-0270(97)02263-2

[21] Aleksander Sadikov, Jure Žabkar, Martin Možina, Vida Groznik, Dag Nyholm, and Mevludin Memedi. 2015. Feasibility of Spirography Features for Objective Assessment of Motor Symptoms in Parkinson's Disease. In *Artificial Intelligence in Medicine*, Lucia Sacchi John H Holmes, Riccardo Bellazzi and Niels Peek (Eds.). Springer International Publishing, Cham, 267–276.

[22] Arash Salarian, Heike Russmann, Christian Wider, Pierre R Burkhard, Françios J G Vingerhoets, and Kamiar Aminian. 2007. Quantification of Tremor and Bradykinesia in Parkinson's Disease Using a Novel Ambulatory Monitoring System. *IEEE Transactions on Biomedical Engineering* 54, 2 (2007), 313–322. https://doi.org/10.1109/TBME.2006.886670

[23] Claudia Schulte and Thomas Gasser. 2011. Genetic basis of Parkinson's disease: inheritance, penetrance, and expression. *The Application of Clinical Genetics* 4 (2011), 67–80. https://doi.org/10.2147/TACG.S11639

[24] scikit learn. 2024. fclassif. https://parkinsonseurope.org/facts-and-figures/statistics/ Accessed: 7-3-2024.

[25] scikit learn. 2024. mutual info classif. https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.mutual_info_classif.html#sklearn.feature_selection.mutual_info_classif Accessed: 7-3-2024.

[26] scikit learn. 2024. Select K Best. https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectKBest.html Dostopano: 28-02-2024.

[27] Hyeongseok Son, Junyong Lee, Jonghyeop Lee, Sunghyun Cho, and Seungyong Lee. 2021. Recurrent Video Deblurring with Blur-Invariant Motion Estimation and Pixel Volumes. *ACM Trans. Graph.* 40, 5, Article 185 (2021), 18 pages. https://doi.org/10.1145/3453720

[28] Molly M Sturman, David E Vaillancourt, , and Daniel M Corcos. 2005. Effects of aging on the regularity of physiological tremor. *Journal of neurophysiology* 93, 6 (2005), 3064–3074.

[29] Stefan Williams, Samuel D Relton, Hui Fang, Jane Alty, Rami Qahwaji, Christopher D Graham, and David C Wong. 2020. Supervised classification of bradykinesia in Parkinson's disease from smartphone videos. *Artificial Intelligence in Medicine* 110 (2020), 101966. https://doi.org/10.1016/j.artmed.2020.101966

[30] Tianze Yu, Kye W Park, Martin J McKeown, and Jane Z Wang. 2023. Clinically Informed Automated Assessment of Finger Tapping Videos in Parkinsonrsquo;s Disease. *Sensors* 23, 22 (2023). https://doi.org/10.3390/s23229149