

# Using Machine Learning to Evaluate Handwriting Patterns for the Diagnosis of Parkinson's Disease

Daniel Roudnitsky; Mahir Sudad; Shlok Khandelwal; Amir Nili; Vlad Dobrin; Aishwarya Jayan; Gashaye Furo; Asbah Shafi; Allegra White; Josselyn Marroquin; Malath Al Matouq; Zoe Roussos; Noriana Jakopin; Nerrise Favour Njunkeng

## Abstract

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that is difficult to diagnose by objective measures, much more so in the early stages. One of the most significant early symptoms is micrographia, which features small handwriting and difficulty maintaining scale or amplitude of movement. Such hypokinetic manifestations of the disease make complex sequential movements difficult.

This study uses various machine learning algorithms such as Support Vector Machines (SVMs), Voting Classification by combining Stochastic Gradient Descent and SVM, an Adaboost model, a Multilayer Perceptron model on the NewHandPD dataset. Transfer learning and Google Inception V3 also implemented image classification. The dataset consists of handwritten exams of a healthy group of 35, and a patient group of 31. It comprises images of spirals, meanders, and circles drawn by the two groups and the signals during the diadochokineses test. Essential features such as x-y coordinates, the pen pressure, and the grip angle are included in the dataset.

Image classification, which was based on the spiral drawings of patients, attained the highest testing accuracy of 87.5%. A central goal of our research efforts is to identify the best methods for testing patients suspected to have Parkinson's and using a modular set of machine learning algorithms of the highest accuracy as a critical part of real diagnostic efforts. The current investigation is one of many attempts at creating a highly predictive assembly of PD diagnostic protocols.

## Introduction

"The hand failing to answer with exactness to the dictates of the will" - James Parkinson

The prevalence of PD is 0.2% in western countries with age of onset generally between 50 and 65 years. One significant symptom PD patients experience is a dramatic change in handwriting known as micrographia (abnormally reduced letter size). PD also affects kinematic features (e.g., speed, acceleration, and stroke duration) of handwriting movements (Ebersbach et al. 1999; Lacquaniti et al. 1987; Van Gemmert et al. 1999; Weiss et al. 1996). Weiss et al. (1996) found that speed scaling was preserved for small elbow-movement amplitudes but not larger ones, leading to the suggestion that a reduced capability to initiate and regulate force causes observed slowness in patients with PD. Similarly, it was suggested that when scaling stroke size in a handwriting task, PD patients had problems controlling force amplitude (Van Gemmert et al. 1999). Ebersbach et al. (1999) inferred from diminished stride length in PD patients that PD induced reductions in force gain. They proposed that PD patients do not have difficulty with force development but rather with maintaining a constant amount of force. The handwriting alterations in patients appear years before they are clinically diagnosed. Hence, handwriting can be used as a significant biomarker for diagnosing PD.

According to Thomas et al. (2017), there are two types of micrographia:

1. Consistent micrographia: is when the handwriting size is reduced throughout the writing. It is related to dysfunction of the basal ganglia motor circuits, possibly because of dopamine depletion.
2. Progressive micrographia: on the other hand is when handwriting size decreases as one continues writing. It is thought to be caused by disconnections between the anterior supplementary motor area, the rostral cingulate motor area, and the cerebellum.

While dopa-mergic drugs bring significant improvement to consistent micrographia, there is no substantial improvement seen in progressive micrographia. This implies that the latter is indicative of late-stage PD (Thomas et al. 2017). Furthermore, patients who had severe micrographia (defined as the demonstration of a reduction in writing surface area >50% on a writing task) showed even higher correlations with measures of disease severity and cognitive impairment (Shukla et al. 2012).

Several studies have shown that PD patients display deficits in bimanual coordination/diadochokinesis due to pathologies in the thalamus and basal ganglia, generally resulting in inter-limb coupling (Daneault et al. 2015). Thus, the diadochokineses test conducted in this dataset consists of individuals holding a pen with straight arms and performing hand-wrist movements. Signals were only generated from these movements; no drawings were involved.

In the NewHandPD dataset, the features extracted from the images and signals classify individuals into Parkinson's vs. Non-Parkinson's groups using different machine learning algorithms, such as Support Vector Machines (SVMs), Voting Classification by combining Stochastic Gradient Descent and SVM, an Adaboost model, and a Multilayer Perceptron model. Transfer learning and Google Inception V3 also implemented image classification. The highest accuracy was obtained by applying image classification with an 87.5% testing accuracy. Image Classification was based on the spiral drawings of the patients. Thus, it establishes an accurate and relatively simple method for Parkinson's disease diagnostics. A potential suspect of PD needs to draw a spiral, and the classifier would determine whether they have Parkinson's with high accuracy.

## Resources

Botucatu Medical School, São Paulo State University - Brazil  
HandPD and NewHandPD datasets

## Methods and Materials

The newHandPD handwriting data set provided by the Botucatu Medical School in São Paulo State University - Brazil was analyzed in this project. The NewHandPD dataset consists of the spiral, meander, and circle drawings 66 individuals divided into 35 control subjects (Healthy Group) and 31 Parkinson's patients (Patient Group). It also includes the participants' movement signals during the Diadochokineses exam.

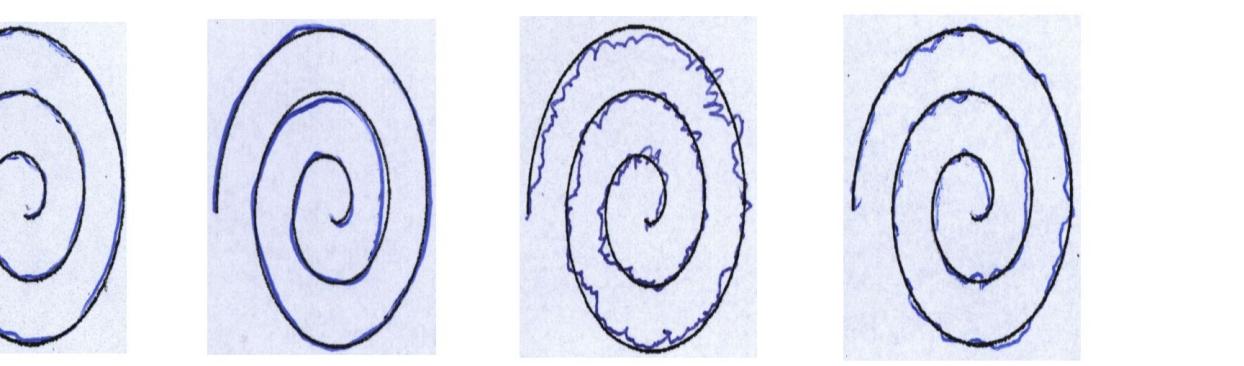


Figure 1. Some examples of spirals extracted from newHandPD dataset: (a) 58-years old male and (b) 28-years old female individuals of control group, and (c) 56-years old male and (d) 65-years old female individuals of patient group.

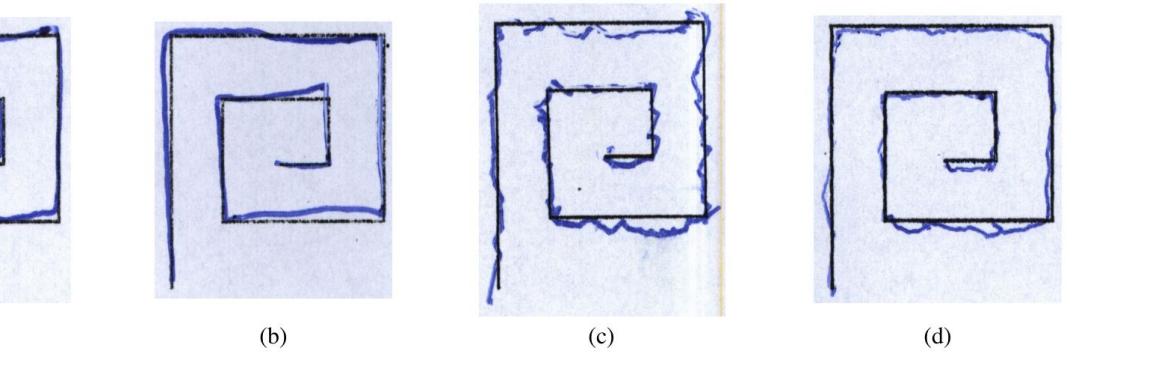


Figure 2. Some examples of meanders extracted from newHandPD dataset: (a) 58-years old male and (b) 28-years old female individuals of control group, and (c) 56-years old male and (d) 65-years old female individuals of patient group.

Several Supervised Learning methods were carried out to analyze the datasets.

- Using the combined datasets, supervised learning models including XGBoosting, Adaboost, Stochastic Gradient Boosting, Support Vector Machines, and Random Forest Trees were created. Later on, an ensemble paradigm was implemented using a Voting Classifier to compare the created models.
- The Inception V3 Convolutional Neural Networks modules were also implemented. The model was then transferred to the NewHandPD data set by using Transfer Learning. The 264 spirals, consisting of 140 healthy and 124 Parkinson's subjects' drawings, were fed to the image classifier. This algorithm avoids classification based on gender, age, or hand-dominance of the subjects.
- A Support Vector Machine classifier was implemented with K-fold cross-validation of 10 folds. The features of the spiral and Meander drawings were combined and fed to the classifier.
- A Multi-Layer Perceptron Classifier with K-fold cross-validation of 10 folds was implemented by using the spiral drawings.
- The signals were featurized using continuous wavelet transforms and classified using the Adaboost ensemble model.

## Results

- The SVM model had little to no false positives, but many false negatives unlike all the boosting methods such as XGBoost, Adaboost, and Stochastic Gradient Boosting. After combining the Stochastic Gradient Boosting technique which had an accuracy of 84.85% and the SVM Model with an accuracy of 84.85% a voting classifier of both of them resulted in an 86.36% accuracy with a 50% split for testing data.
- After running a Support Vector Machine Classifier with 10-fold cross validation, an accuracy of 84.3%, and a Matthew Correlation coefficient of 0.78 was achieved with the 'RBF' kernel. Dimensionality was reduced with Isometric mapping.
- After running a Multi-layer Perceptron classifier with the 'lbfgs' solver with 10 fold cross-validation, an evaluation accuracy of 84.3% was achieved with a Matthew Correlation Coefficient of 0.49. Dimensionality was reduced with Principal Component Analysis.
- Another approach involved the feature extraction of the continuous wavelet transform (CWT) of the subjects' signals while taking the Diadochokineses test. Using the Ricker wavelet as the mother wavelet, the CWT of the signal was taken, and features were extracted from the CWT by finding the peaks of the CWT by smoothing it and considering the relative maxima which appear at enough length scales. AdaBoost was trained on these extracted features and achieved 100/92.3/71.42 percent accuracy across train/test/validation with a 60/20/20 split, implying significant overfitting and poor generalization with this approach.
- Another model was designed based on computer vision techniques. This method involved training Google's Inception V3 model and then using transfer learning to retrain with the spiral handwritings. This method resulted in the highest test accuracy among the other machine learning algorithms. It resulted in a 100% learning accuracy, 94.01% validation accuracy, 87.5% test accuracy, and a cross entropy value of 0.002. These results were obtained by splitting the data set into an 80% learning set, a 10% test set, and a 10% validation set.

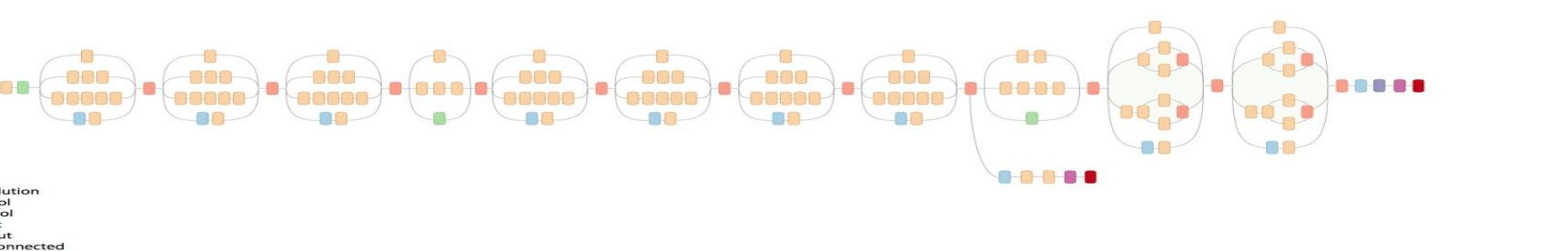


Figure 3. Visual representation of the Google Inception V3 model architecture.

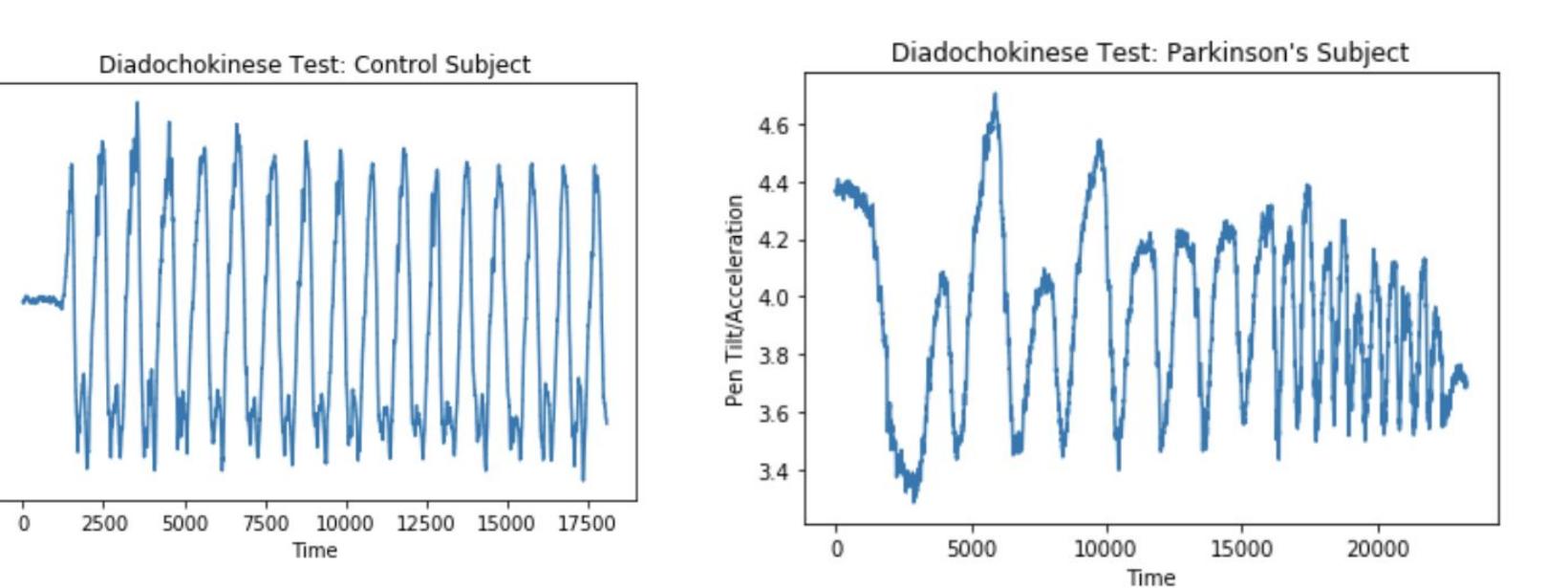
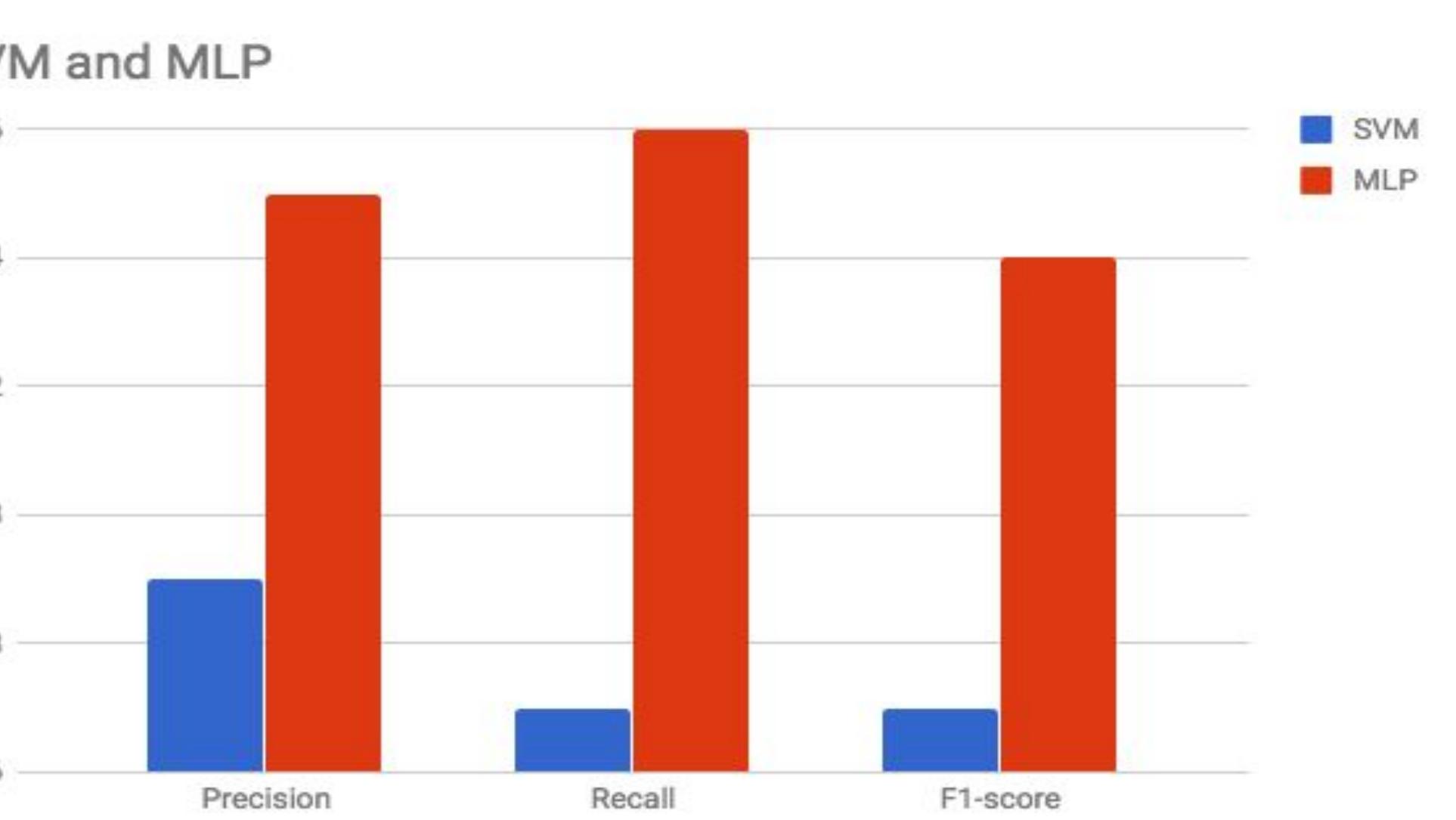


Figure 4. Comparison of the Pen Tiltacceleration between Parkinson's and healthy groups.

Table 1. Results of the SVM and MLP Classifiers

	Precision	Recall	F1-score	Support
SVM	0.79	0.77	0.78	66
MLP	0.85	0.86	0.84	49

Chart 1. Histogram Representation of the results of SVM and MLP



## References

1. Daneault, J. F., Carignan, B., Sadikot, A. F., & Duval, C. (2015, August). Inter-limb coupling during diadochokinesia in Parkinson's and Huntington's disease. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4540003/>
2. Ebersbach G et al. Comparative analysis of gait in Parkinson's disease, cerebellar atrophy and subcortical arteriosclerotic encephalopathy. Brain 1999 Jul;122:1349-1355
3. Lacquaniti F, Ferrigno G, Pedotti A, Soechting JF, Terzuolo C (1987) Changes in spatial scale in drawing and handwriting: Kinematic contributions by proximal and distal joints. J Neurosci 7:819-828
4. Margolin, D.I., Wing, A.M. Graphia and micrographia: clinical manifestations of motor programming and performance disorders. Acta Psychol. 54 (1983), pp. 263-283.
5. P. Drotar, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, 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 u.
6. Pereira, C. R., Weber, S. T., Hook, C., & Ross, G. H. (2016). Deep Learning-aided Parkinson's Disease Diagnosis from Handwritten Dynamics, 1-7. Retrieved from <https://pdfs.semanticscholar.org/7988/3c397222b8cb59f608df2f8889dec15bf5.pdf>
7. Shukla, A. W., Ounpraseuth, S., Okun, M. S., Gray, V., Schwankhaus, J., & Metzger, W. S. (2012). Micrographia and related deficits in Parkinson's disease: A cross-sectional study. BMJ Open, 2(3), doi:10.1136/bmjopen-2011-000628
8. Thomas, Mathew & Lenka, Abhishek & Pal, Pramod. (2017). Handwriting Analysis in Parkinson's Disease: Current Status and Future Directions. Movement Disorders Clinical Practice. 4. 10.1002/mdc.21552.
9. Van Gemmert AW, Teulings HL, Contreras-Vidal JL, Stelmach GE. Parkinson's disease and the control of size and speed in handwriting. Neuropsychologia. 1999; 37(6):685-694
10. Weiss P, Stelmach GE, Adler CH, Waterman C (1996) Parkinsonian arm movements as altered by task difficulty. Parkinsonism Relat Disord 2:215-222

## Limitations of the Study

- In regards to processing the movement signals, one of the extracted features was age, which could introduce bias as the model may have learned that the older a person is, the more likely they have Parkinson's. However, when Age was removed from the model, accuracy dramatically reduced to approximately 60%; this is why age was left in the extracted feature set. Next, an attempt was made to create models for each of the spirals and meander images separately. However, the accuracies for the model were poor (approximately around 70%), so for each patient, the meander and spiral extracted features were combined into one dataset. The goal of this was to provide the model with more information about each patient so that it could better learn the differences between healthy individuals and patients.
- In the Image Classification model, learning based on the meander drawings was avoided, because it resulted in lower accuracy rates, compared to learning on the spiral data set. Combining the two drawing sets, resulting in a slightly higher accuracy (88.9%), however, the validation accuracy (78%), indicated that the model was overfitting. Hence, the Image Classification was solely trained on the spiral drawings.
- Although there were different types of data such as images and signals in the dataset, the sample size was not large enough. There were only a total of 264 spiral images, and 267 meander images, and 372 signals from patients. Smaller sample sizes usually correspond to lower evaluation accuracies, and worse results compared to larger sample sizes.

## Conclusions

Diagnostic tests that exploit the early hypokinetic manifestations of PD, such as micrographia, have been shown to be adequate in the differential diagnosis of early Parkinson's. Using the NewHandPD dataset, different supervised machine learning algorithms were conducted, and results were compared. It was concluded that Image Classification by performing transfer learning with Google's Inception V3 model, in particular, leads to higher accuracy rates. This model was trained on 264 spiral images, which resulted in 87.5% testing, 94.01% validation, and 100% learning accuracy.

Parkinson's is one of the hardest neurodegenerative diseases to objectively diagnose, especially in early stage and, at times, even in the middle of the disease progression. There are many reasons why this is the case, not the least of which is its significant overlap with other movement disorders. It is believed that cognitive manifestations of PD occur in later stages, even though some evidence points to the contrary.

We have previously obtained good clinically acceptable results applying machine learning to audio data of Parkinsonian patients. This study also shows the potential of more refined protocols to test the kinematic features of handwriting in patients suspected of having Parkinson's. Tests that better capture the speed, acceleration, and stroke duration of handwriting movements would hopefully allow for better feature extraction/engineering methodologies.

## Future Directions

In the future, we would like to analyze a larger dataset, to provide better evaluation accuracy. A larger dataset is more likely to ensure that the learning algorithm reaches 97.5% accuracy, per medical standards. We would also like to build a more robust convolutional neural network model for classification that is tuned explicitly towards the dataset, instead of retraining with transfer learning on a different model. This model will include a high cost for misdetection and allow cost for a false alarm.

Moreover, accessing raw data from various clinical protocols would allow us to identify an ultimate set of features which should be extracted from any patient exhibiting Parkinsonian-like symptoms. This would further inform the design of a package of objective diagnostic tools that would finally eliminate the human decision-making errors among practicing clinicians.

## Acknowledgements

Dr. Nevine El-Leithy, Deep Brain Neurotechnologies Research Director  
Dr. Robert Newcomb, Department of Electrical and Computer Engineering  
Food and Drug Administration (FDA)  
Deep Brain Neurotech FIRE Stream  
First-Year Innovation & Research Experience (FIRE)

- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5585598/>
- PD is a disorder of the central nervous system that has an effect on controlling muscles; thus, it influences movement, speech, and handwriting of patients (Pahwa, Lyons, and Koller, 2007). The handwriting of a PD patient (PDP) is often characterized by micrographia, a reduction of letter size during continuous writing (Teulings and Stelmach, 1991). It has also been reported that kinematic features (e.g., speed, acceleration, and stroke duration) of handwriting movements are affected by PD (Teulings and Stelmach, 1991; Tucha *et al.*, 2006; Flash *et al.*, 1997; Longstaff *et al.*, 2001). (<https://www.sciencedirect.com/topics/medicine-and-dentistry/handwriting>)
- Because studies based on kinematic analyses of handwriting have reported specific abnormalities in different PD phenotypes and in several PD mimics, such as ET and PSP, these may be used as clinical markers during the early stages of disease. Because PD is sensitive to medications, handwriting may also be a reliable marker of disease progression.  
(<https://onlinelibrary-wiley-com.proxy-um.researchport.umd.edu/doi/epdf/10.1002/mdc3.12552>)
- The Parkinson's Disease Handwriting Database (PaHaW) consists of multiple handwriting samples from 37 parkinsonian patients (19 men/18 women) and 38 gender and age matched controls (20 men/18 women). The database was acquired in cooperation with the Movement Disorders Center at the First Department of Neurology, Masaryk University and St. Anne's University Hospital in Brno, Czech Republic. Each subject was asked to complete a handwriting task according to the prepared filled template at a comfortable speed. The completed template was shown to the subjects; no restrictions about the number of repetitions of syllables/words in tasks or their height were given.
- A tablet was overlaid with a empty paper template (containing only printed lines and square box specifying area for Archimedean spiral), and a conventional ink pen was held in a normal fashion, allowing for immediate full visual feedback. The signals were recorded using the Intuos 4M (Wacom technology) digitizing tablet with 200 Hz sampling frequency.
- Digitized signals were acquired during the movements executed while exerting pressure on the writing surface and during the movement above the writing surface. We denote these signals as on-surface movement and in-air movement, respectively. The perpendicular pressure exerted on the tablet surface was also recorded. The recordings started when the pen touched the surface of the digitizer and finished when the task was completed. the tablet captured the following dynamic features (time-sequences): x-coordinate; y-coordinate; time stamp; button status; pressure; tilt; and elevation. Button status is a binary variable, being 0 for pen-up state (in-air movement) and 1 for pen-down state (on-surface movement).
- SVC file structure
- 1st line: number of samples
- n-th line: Y coordinate, X coordinate, time stamp, button state, azimuth, altitude, pressure
- File corpus\_PAHAW.xlsx provides subject characteristics.
- For more information about data see:
- Drotar P, Mekyska J, Rektorova I, Masarova L, Smekal Z, Faundez-Zanuy M. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artif Intell Med.* 2016; XX(XX): XX–XX.
- Drotar P, Mekyska J, Rektorova I, Masarova L, Smekal Z, Faundez-Zanuy M. Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease. *Comput Methods Programs Biomed.* 2014; 117(3): 405–411. (info from source of the data)

<https://onlinelibrary-wiley-com.proxy-um.researchport.umd.edu/doi/epdf/10.1002/mdc3.12552>

<https://www.sciencedirect.com.proxy-um.researchport.umd.edu/science/article/pii/S0933365716000063?via%3Dihub#sec0005>

<https://link.springer.com.proxy-um.researchport.umd.edu/content/pdf/10.1007%2Fs00415-013-6996-x.pdf>

<https://link.springer.com/article/10.1007%2Fs00221-009-1925-z>

## Outline

- intro to PD
- neuroscientific info about handwriting changes due to PD
- changes in handwriting patterns in early vs late stage PD
  - “Alterations in the kinematics of handwriting are among the recently proposed biomarkers of PD”
  - “two prominent handwriting features observed in the group of PD patients were progressively decreasing size of letters and the ascending direction of the phrase with respect to the horizontal.” - Handwriting Rehabilitation in Parkinson Disease: A Pilot Study
  - “Because PD is sensitive to medications, handwriting may also be a reliable marker of disease progression.” Handwriting Analysis in Parkinson's Disease: Current Status and Future Directions
  - 
  -
- is it just tremor or eye-hand coordination?????

## methods

- source of our data
- how it was collected
- what is done

Gallicchio et al proposed an approach for diagnosing PD based on Deep Echo State Networks (ESNs) by analyzing whole time-series collected from a tablet device during the sketching of spiral tests. Their results showed that the predictive accuracy obtained from DeepESN outperforms standard shallow ESNs in training, validation and test set of 2.67%, 2.95%, and 3.07% of accuracy respectively.

Pereira et al introduced Restricted Boltzmann Machines (RBMs) as a method for automatic identification of PD by means of images acquired from handwritten exams. Their results showed that for their Meander 64 x 64 dataset, there was a 74.38% accuracy and for their spiral 64 x 64 dataset 83% accuracy. To classify samples as PD or controls,