NeuroTrace: A Novel, Machine-Powered Software to Detect Neurodegeneration through Handwriting Kinematics Analysis

AzSEF 2024

SR-TMED-003

Introduction

Neurodegenerative diseases (NDs), such as Parkinson's and Alzheimer's affect more than 62 million people worldwide [1].

- Occurring due to progressive degeneration of nerve cells, NDs can be correlated to lifestyle, diet, genetics, injury, or even exposures to hazards in a workplace
- Most NDs, like Parkinson's, share a common pattern of **damaging the motor system**
  - One motor degeneration is the loss of **hand control.**
    - Hand motion **can** become **segmented and jagged.**
    - **Daily activities,** like **writing, are impeded.**

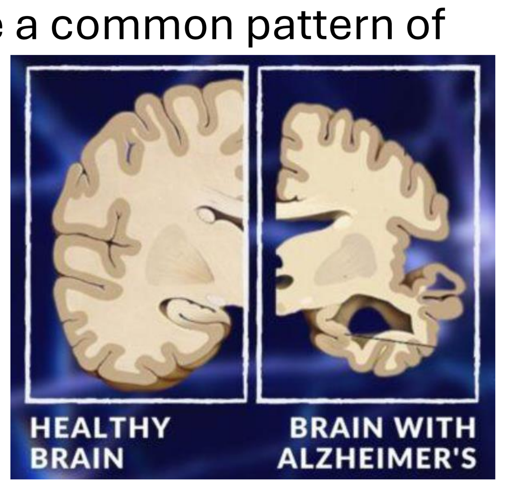


Figure 0 Credits: biologyonline.com

Previous Research

<u>Digitizing Tablets and Pens</u>, along with <u>handwriting kinematics (HKs)</u>, have been utilized in previous studies:

- Mergl et. al [2] utilized a tablet to analyze HKs & drawing kinematics **to correlate depression and mental disorders** with handwriting.
- Isenkul et. al [3] tested for **Parkinson's** and add to Kaggle dataset using WACOM tablet
  - Yet, their **experiment only used the Archimedes Spiral task,** limiting full comprehensive understanding of NDs.
- Nachum et. al [4] designed a similar proj. involving computer vision, diagnosing Alzheimer's & Parkinson's.
  - However, they **utilized camera tracking instead,** and **were unable to obtain participant samples.**

Figure 1 & 2 Credits: Isenkul et. al [4]

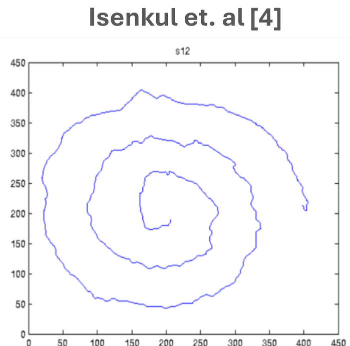


Figure 1: control subject's spiral drawing

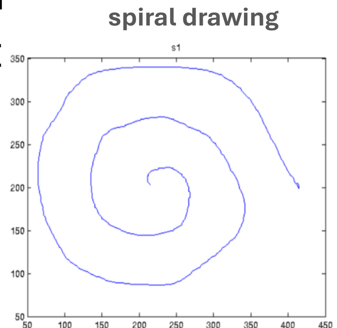


Figure 2: Parkinson's subject's spiral drawing

Problem

Current methods for NDs detection involve **manual assessments,** which cause **slow, ineffective,** and **universally inapplicable** NDs detection & diagnoses.

Goal

Develop a **portable, automated, quick, efficient,** and **accurate** NDs detection system involving handwriting kinematics with current technology like AI

Process Outline

Preprocess DARWIN database → Train RandomForest Classifier → Tune, adjust, & finalize model → Design & code NeuroTrace test frontend → Gather participants for prototype testing → Train/Test on subject data & verify software accuracy

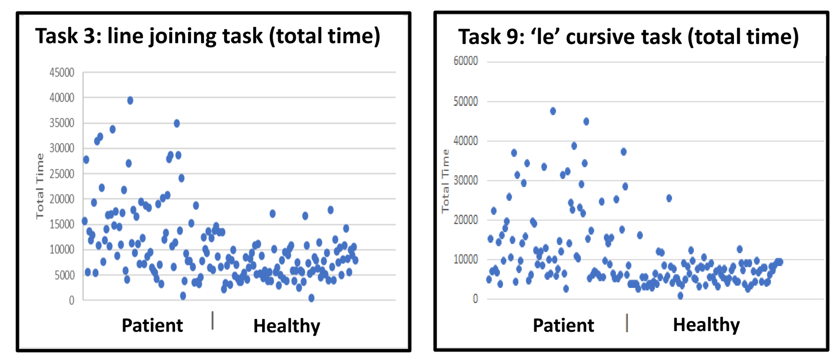
Methods

1. Data Processing/Visualization

NeuroTrace model based on **DARWIN dataset** (n=174), the **most extensive HKs database** for NDs research [5].

- DARWIN participants completed 25 pen tasks, produced 18 metrics
  - **Avg speed, acc, jerk, pressure, etc.**
- NeuroTrace simplified DARWIN down to **6 tasks and 12 metrics**

DARWIN dataset visualized → remove unnecessary tasks/features



Figures 3 & 4: visualizations of subjects' total time taken to perform tasks.

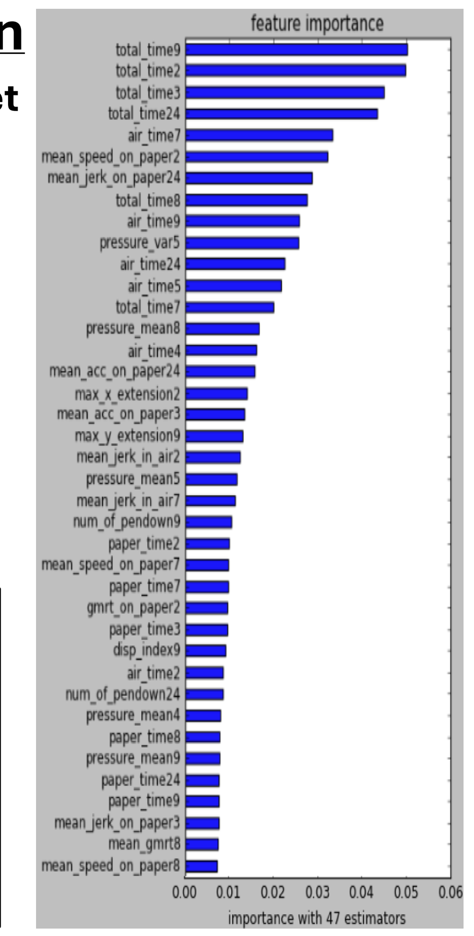


Figure 5: DARWIN metric importance

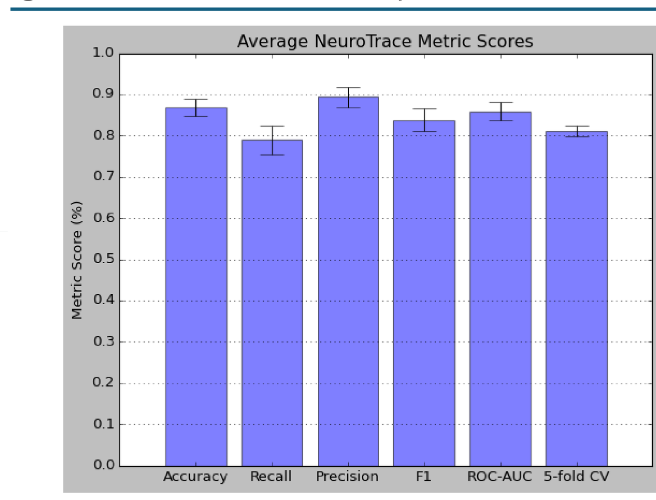


Figure 9: Average performance metrics for trained NeuroTrace model

Figure 10: Confusion matrix for trained NeuroTrace model

2. Model Training & Optimization

- Model trained w/ *scikit-learn*'s **Random Forest Classifier** (Python)
  - RF Classifier **generalizes well to new datasets** (less overfit)
- Outputs prediction (healthy/patient) and analyzable metrics

```
# Splitting data into training and testing
X_train, X_test, Y_train, Y_test =(
  train_test_split(X,Y,train_size=.75))

# Creating Random Forest model
rf = RandomForestClassifier(n_estimators=100)
rf.fit(X_train,Y_train)
Y_pred = rf.predict(X_test)
```

Figure 6: Code for RF test/train

RF Classifier Prototype
- Outputs class (H/P)
- Output analyzable metrics

→ **Output (metrics)**
- Accuracy %
- F1 Score
- Precision
- Recall
- Cross-Validation
- ROC Area-under-Curve

→ **Hyperparameter Tuning**

Adjust:
- Tree Depth
- # Estimators
- Min/Max Split
- Min/Max Nodes
- Random Seed

Hyperparameter Tuning Process

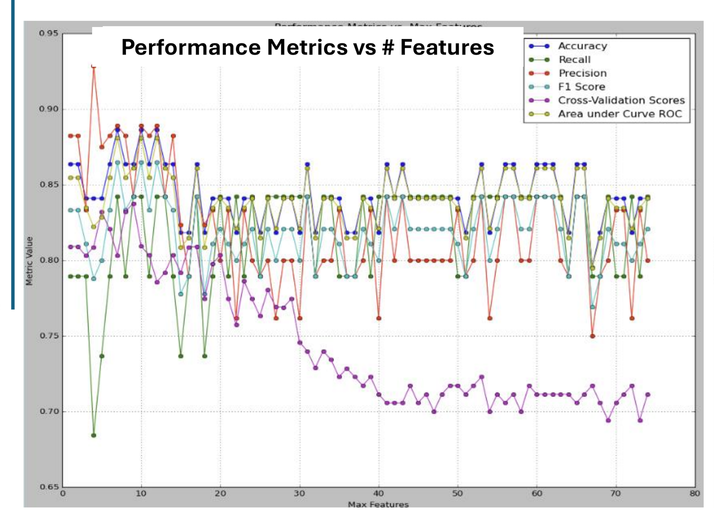


Figure 7: Performance metrics vs # of features **(0-80).** 5-fold CV Score highlighted & drop after certain depth

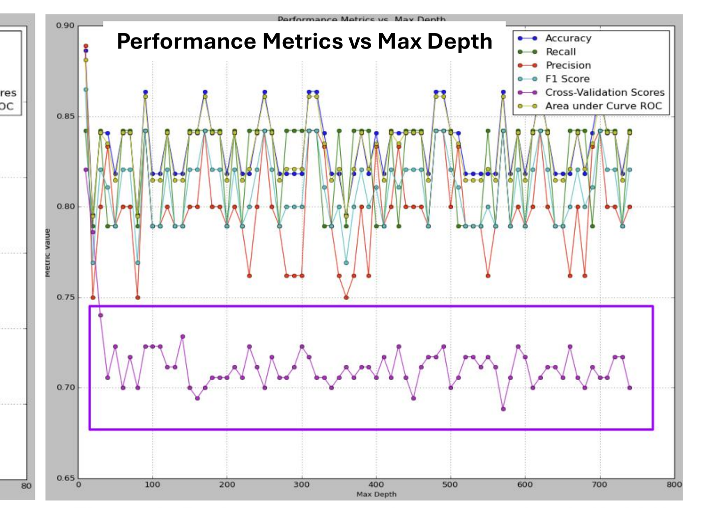


Figure 8: Performance metrics vs max depth **(0-800).** 5-fold CV Score highlighted & drop after certain depth

- Figure 9 indicates that model has **high accuracy (90%)** and **precision (94%),** with **fairly good recall (84%).** The **F1 score** (weighted avg. of precision/recall) **is 88%.**
  - <u>Intuitively, the model is **very reliable** at identifying/flagging cases (some false +s)</u>
- The **5-fold Cross-validation score (80%)** and **ROC Area Under Curve score (90%)** indicate **<u>NeuroTrace generalizes well across unknown datasets</u>** <u>(real-life data).</u>
- This data shows **NeuroTrace results are statistically significant**, yielding **significantly better results than random chance (~50%).**
- In Figure 9, the **error bars do not overlap with 0.5,** indicating statistical significance.

3. NeuroTrace Frontend Development

- Self-developed **portable & interactive HTML program utilizing JavaScript**
- **WACOM Intuos Tablet collects pen data** w/o actual ink, testing hand-eye coordination
- Interactive, simple app with 6 tracing tasks
  - **Captures pen kinematics** (x,y, pressure, tilt, pen ups/downs) from screen-mapped **WACOM tablet** & **exports CSV file**
- CSV file **processed & normalized** to DARWIN metrics **→ analyzed by RF model**

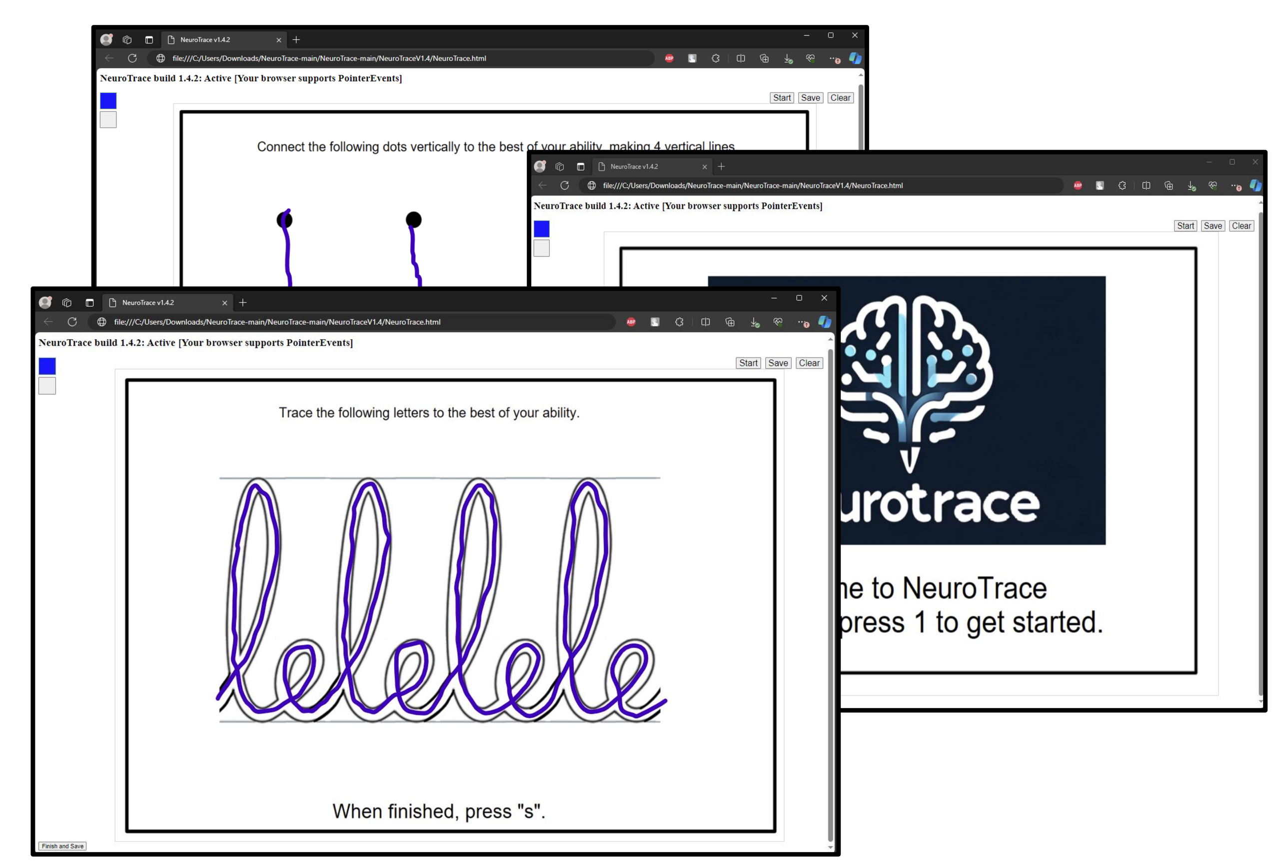
Figure 11: Data collection device - Wacom Intuos Tablet (120hz) **Credit: pbtech.com**

$$\text{avg. speed} = \frac{1}{n}\sum_{i=0}^{n}\sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}$$

$$\text{avg. acceleration} = \frac{1}{n}\sum_{i=0}^{n}\frac{\sqrt{(x_{i+1} - 2x_i + x_{i-1})^2 + (y_{i+1} - 2y_i + y_{i-1})^2}}{120}$$

$$\text{avg. jerk} = \frac{1}{n}\sum_{i=0}^{n}\frac{\sqrt{(x_{i+2} - 3x_{i+1} + 3x_i - x_{i-1})^2 + (y_{i+2} - 3y_{i+1} + 3y_i - y_{i-1})^2}}{120^2}$$

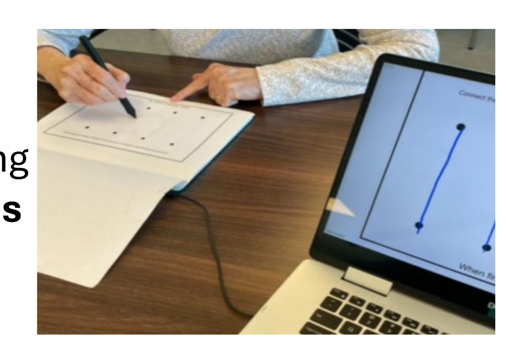
Figure 12: Central difference approximation formulas for parametric/kinematics equations, which calculate for the tested handwriting kinematics



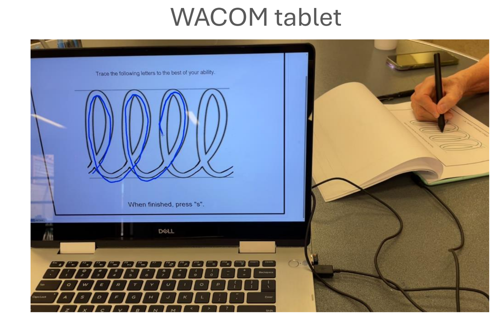
Figures 13-15: Performance metrics vs # of features (0-80). 5-fold CV Score highlighted & drop after certain depth

4. Real-Life NeuroTrace Prototype Testing

- **Randomly recruited seniors** (n=18, mean age=80) from nearby senior facilities participated in NeuroTrace pilot testing
  - **All subjects viewed and signed informed consent forms** (explained purpose of study and data protection policies)
- Subjects completed two sets of tasks:
  1. Completed a quick **demographics survey** about subjects' **history of NDs** and **educational & vocational history**
     - **Survey determine if subjects were positive (Patient) or negative (Control) baseline for data collection**
  2. Utilized the (set up) **WACOM tablet and digital pen** to complete the **6 NeuroTrace tracing tasks**
- **Entire testing process took < 10 minutes on average**



Figures 16 & 17: Participant completing vertical dots' task (above) & the cursive 'l' task (below) w/ WACOM tablet



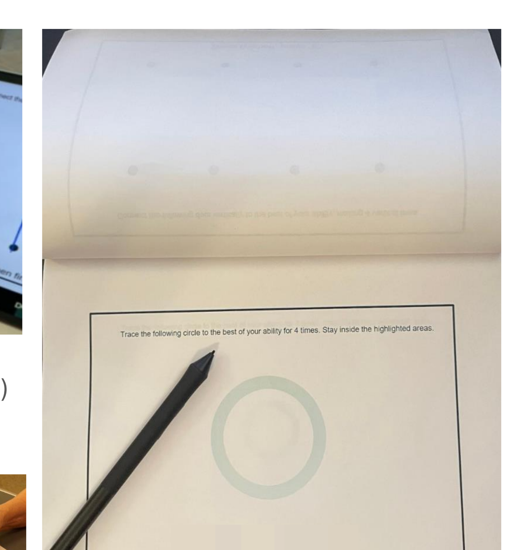


Figure 18: Data collection WACOM tablet, set up w/ tracing assignment & ready for participant tracing

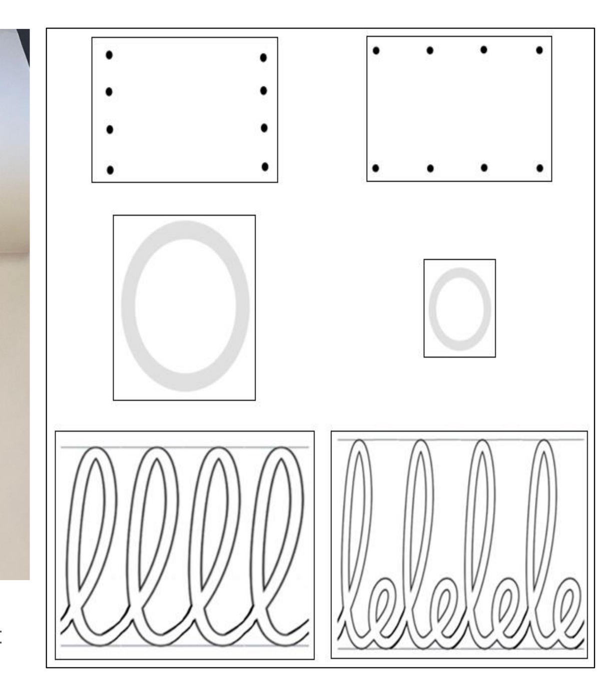


Figure 19: Each of the 6 NeuroTrace tasks which subjects were instructed to trace out

All figures created by student unless otherwise specified

Data Analysis

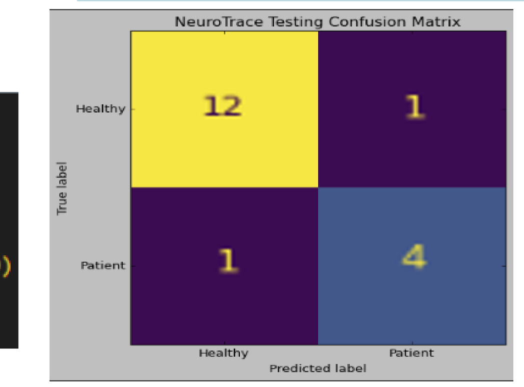


Figure 20: Confusion matrix for NeuroTrace participant testing

	Accuracy	Precision	Recall	F1 Score
Training	90%	94%	84%	88%
Senior Testing	88%	80%	80%	80%

Figure 21: NeuroTrace training and real-life testing accuracy metrics

- Figures 20 & 21 indicate that the model has **high accuracy (88%) and decent precision/recall (80%)** with **5-fold cross-validation** when analyzing the collected data.
  - <u>Thus, NeuroTrace is **reliable** at identifying NDs cases based on the tested seniors' kinematic data (n=18).</u>
- The results prove that a **distinct difference between healthy control subjects' and NDs subjects' handwriting kinematics exists** (which the trained Random Forest Classifier identified)

*Yet, it is important to note that these statistics are based on a **limited sample size***

Conclusion

<u>Based on real-life participant data, NeuroTrace will:</u>

- **Ensure early NDs detection** with an **accuracy > 80%**
- Contribute to handwriting **and kinematics relevance** in the field of **neurodegenerative disease research**

<u>NeuroTrace has several implications and applications:</u>

- **Quick, accurate,** and **non-invasive screening tool**
- **Cheap, efficient** application for **low-resource groups**
- Effective **remote diagnostic** and **medical monitor**
- Useful **longitudinal study tool,** quick & simple tasks

<u>Ultimately, NeuroTrace creates a future where the burden of neurodegenerative diseases is alleviated, increasing quality-of-life for those who need and deserve it most.</u>

Future Research

- **Multimodal software integration,** such as **image analysis or other sensory data** (speech, hearing, etc.) will provide more accuracy & model complexity
- **Utilize more datasets to enhance model depth**
- **Personalization implementation** for **users' individual profiles** to improve accuracy:
  - Demographics (age, ethnicity)
  - Genetic history
  - Vocational/educational history
  - Lifestyle (diet, hobbies, health)

Fig. 22: Multimodal software integration **Credits: brainsight.com**

References

1.News-Medical. "Alzheimer's Disease | Definition, Causes, Diagnosis & Treatment." News-Medical.net, 25 Oct. 2019, www.news-medical.net/health/Alzheimers-Disease-Definition-Causes-Diagnosis-Treatment.aspx.

2.Mergl, R., et al. "Kinematical Analysis of Handwriting Movements in Depressed Patients." Acta Psychiatrica Scandinavica, vol. 109, no. 5, May 2004, pp. 383–391, https://doi.org/10.1046/j.1600-0447.2003.00262.x.

3.Isenkul, Muhammed, et al. "Improved Spiral Test Using Digitized Graphics Tablet for Monitoring Parkinson's Disease." The 2nd International Conference on E-Health and Telemedicine , vol. pp: 171-175, May 2014, https://doi.org/10.13140/RG.2.1.1898.6005.

4.Nachum, Ron, et al. "A Novel Computer Vision Approach to Kinematic Analysis of Handwriting with Implications for Assessing Neurodegenerative Diseases." IEEE Xplore, 1 Nov. 2021, ieeexplore.ieee.org/document/9630492.

5.Cilia, Nicole, et al. "DARWIN Dataset." UCI Machine Learning Repository, 5 Nov. 2018, https://archive.ics.uci.edu/dataset/732/darwin.