IEEE Journal of Translational Engineering in Health and Medicine

Received 22 August 2023; revised 17 November 2023 and 4 December 2023; accepted 9 January 2024.
Date of publication 18 January 2024; date of current version 6 February 2024.

Digital Object Identifier 10.1109/JTEHM.2024.3355432

NeuroDiag: Software for Automated Diagnosis of Parkinson's Disease Using Handwriting

QUOC CUONG NGO[1], (Member, IEEE), NICOLE MCCONNELL[2], MOHAMMOD ABDUL MOTIN[1], (Senior Member, IEEE), BARBARA POLUS[1], ARUP BHATTACHARYA[2], SANJAY RAGHAV[3], AND DINESH KANT KUMAR[1], (Senior Member, IEEE)

[1]School of Engineering, STEM College, RMIT University, Melbourne, VIC 3000, Australia
[2]Goulburn Valley Health, Shepparton, VIC 3630, Australia
[3]Monash Medical Centre, Department of Neurosciences, Clayton, VIC 3168, Australia

CORRESPONDING AUTHOR: D. K. KUMAR (dinesh.kumar@rmit.edu.au)

This work was supported in part by Jesse Medical Technologies.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Goulburn Valley Health Human Research Ethics Committee under Approval No. HREC/74760/GVH-2021-258233 and the RMIT University Human Research Ethics Committee under Approval No. 2021-24384-14138 and No. 2020-19347-11498.

ABSTRACT Structured Abstract — *Objective*: A change in handwriting is an early sign of Parkinson's disease (PD). However, significant inter-person differences in handwriting make it difficult to identify pathological handwriting, especially in the early stages. This paper reports the testing of NeuroDiag, a software-based medical device, for the automated detection of PD using handwriting patterns. NeuroDiag is designed to direct the user to perform six drawing and writing tasks, and the recordings are then uploaded onto a server for analysis. Kinematic information and pen pressure of handwriting are extracted and used as baseline parameters. NeuroDiag was trained based on 26 PD patients in the early stage of the disease and 26 matching controls. *Methods*: Twenty-three people with PD (PPD) in their early stage of the disease, 25 age-matched healthy controls (AMC), and 7 young healthy controls were recruited for this study. Under the supervision of a consultant neurologist or their nurse, the participants used NeuroDiag. The reports were generated in real-time and tabulated by an independent observer. *Results*: The participants were able to use NeuroDiag without assistance. The handwriting data was successfully uploaded to the server where the report was automatically generated in real-time. There were significant differences in the writing speed between PPD and AMC ($P<0.001$). NeuroDiag showed 86.96% sensitivity and 76.92% specificity in differentiating PPD from those without PD. *Conclusion*: In this work, we tested the reliability of NeuroDiag in differentiating between PPD and AMC for real-time applications. The results show that NeuroDiag has the potential to be used to assist neurologists and for telehealth applications.

INDEX TERMS Automated diagnosis, handwriting, Parkinson's disease, software-based medical devices.
Clinical and Translational Impact Statement — This pre-clinical study shows the feasibility of developing a community-wide screening program for Parkinson's disease using automated handwriting analysis software, NeuroDiag.

I. INTRODUCTION

PARKINSON'S disease (PD) is the second most common neurodegenerative disorder [1] marked by decreased dopamine levels in the brain [2]. It is a multi-symptom disease with complex manifestations of the symptoms [3]. To reduce the likelihood of misdiagnosis, the International Parkinson's and Movement Disorder Society (MDS) has formulated the (UPDRS) guidelines [4]. However, many of the early-stage symptoms such as subtle changes in handwriting and voice [5], can be missed because of the natural variation in the population and changes in an individual with age [6], [7], [8].

Movement disorder symptoms such as tremors, rigidity, bradykinesia, and postural instability are the cardinal

manifestations of PD [9]. Another important early-stage symptom is micrographia [10]. These manifestations significantly affect the dynamics of the sketching and writing ability of people with PD (PPD) causing changes in kinematics [11], the size of writing [10], and pen pressure [12]. Since the cerebral cortex, basal ganglia, and cerebellum are involved in learning and performing handwriting tasks, the assessment of PD can be covered by handwriting and the feature extracted from disrupted handwriting has the potential to be used as a prominent biomarker to identify physiological changes due to PD [8], [13].

The pen-paper-based handwriting analysis is routinely conducted by neurologists when assessing their patients. However, these tasks only allow to analyze the spatial information but do not have the option to investigate the pressure, temporal, or kinematic information. Digital tablets allow the recording and analysis of the dynamics and pressure during writing along with spatial features. This removes the burden of manual inspection of pen-paper-based handwriting and besides micrographia, also indicates other movement symptoms. The correlation between dynamic features during sketching spirals and the severity of the disease was observed in our previous study [14]. Zham et al. [15] have shown that handwriting dynamics are more sensitive and specific for separating PPD and healthy controls. Using handwriting kinematics and pressure, PD was detected with a sensitivity of 87.4% and a specificity of 80.9% [12].

Deep learning models have been used to identify PD based on handwriting patterns. A three-stage fuzzy classifier was used in [16]. While a combination of principal component analysis and deep recurrent neural network model was used in [17], a hybrid model of support vector machine and RESNET-50 was used in [18]. However, it can be observed that in the published deep learning-based PD detection models, the analysis was conducted offline [19], [20], [21] and was only based on the shape of the drawings, while the dynamics were not considered. Handcrafted features of the dynamics of handwriting were proven to outperform the image analysis using deep convolutional neural networks in detecting PD using handwriting data [14], [15], [22].

With the widespread availability of computers and smartphones, a number of software-based medical devices (SBMD) have been developed for diagnosis and monitoring the progression of the disease [23], [24]. The development of such devices aims to reduce clinician bias, reduce the clinician's time for diagnosis, and provide an accurate recording of the symptoms [25].

The review by Linares-del Rey et al. showed that there are more than 100 mobile phone applications that had been designed for providing information, and assisting the assessment and treatment of PD [26]. The techniques have considered different symptoms and modalities such as using the phone accelerometer that was used to collect data about tremors [27], [28], gait [29], [30], and bradykinesia [31]. There is a fairly widespread use of wearable motion sensors which analyse various motor components of the disease. The speech and voice of PPD were assessed using input from the phone microphone [29], [32]. Changes in the dynamics of handwriting were investigated using finger activities recorded from a smartphone multi-touch screen [33]. Nevertheless, while there are a number of SBMD for PD reported in the literature, none of these has yet been reported as being used in clinics.

In this paper, we have experimentally tested NeuroDiag, a software-based medical device for the automated detection of PD using patterns of six handwriting tasks. The front end of this software runs on a Microsoft Surface Pro tablet. It is designed to direct the user through the six tasks and collect the corresponding handwriting data. The data are analysed on the cloud and the result is sent to the user's registered clinic. The proposed system was developed and tested for two different clinics and real-time applications.

II. METHODS AND PROCEDURES

A. PARTICIPANTS

The experimental protocol was approved by Goulburn Valley Health Human Research Ethics Committee (HREC/74760/GVH-2021-258233) and RMIT University Human Research Ethics Committee (2021-24384-14138 and 2020-19347-11498). All participants provided their oral and written consent before the start of the experiment.

People with Parkinson's disease in this study were diagnosed by neurologists with movement disorder speciality using MDS criteria [34]. This was based on the patient's medical history, a review of signs and symptoms, and a motor disorder examination. The severity of the disease was assessed using the Unified Parkinson's Disease Rating Scale (UPDRS-III). PPD were recruited from the two outpatient clinics at i) Goulburn Valley Health, Shepparton, Victoria, Australia, and ii) Dandenong Neurology, Melbourne, Australia.

The control group was recruited to approximately match the age distribution and gender of the PPD from multiple aged care facilities using word-of-mouth and appropriately located posters. We also collected data from young healthy people at RMIT University for reference.

The exclusion criteria were: i) clinically observed or self-reported skeletal injuries, ii) neurological and muscular-skeletal diseases (other than Parkinson's), and iii) excess Levodopa medication causing dyskinesia. The UPDRS-III and severity of PD were observed at the time of the experiment. Dandenong Neurology (DN) is a private clinic and the UPDRS data of participants recruited from this clinic were not available.

In this study, forty-nine people with PD participated; of these, 26 patients were recruited from Goulburn Valley Health, Shepparton, Victoria, Australia (PPD-GVH) and 23 from Dandenong Neurology, Melbourne, Australia (PPD-DN). For the control group, 25 age-matched people participated in the study, seven of whom participated multiple times. The repeated experiments of these participants were

IEEE Journal of Translational Engineering in Health and Medicine

TABLE 1. Demographic and clinical information of participants.

Demographic	PPD-GVH	PPD-DN	AMC-1	AMC-2	YC
Number of Participants	26	23	19	8*	7
Number of records	26	23	26	13	8
Age (years)	72.58 ± 7.21	68.19 ± 11.88	69.21 ± 7.52	64.50 ± 8.94	30.14 ± 3.72
Gender (Male, Female)	15, 11	14, 9	3, 16	3, 3	5, 2
Dominant Hand (right, left)	21, 5	19, 4	17, 2	6, 0	7, 0
Disease Duration (years)	3.91 ± 3.46	3.93 ± 3.33	-	-	-
Unified Parkinson's Disease Rating Scale-III	37.46 ± 17.28	-	-	-	-
Montreal Cognitive Assessment (MoCA)	25.77 ± 3.46	-	-	-	-

*Two age-matched healthy people participated in both the developing and testing phases. PPD-GVH: People with Parkinson's disease recruited from Goulburn Valley Health; PPD-DN: People with Parkinson's disease recruited from Dandenong Neurology; AMC-1: Group 1 of age-matched controls; AMC-2: Group 2 of age-matched controls; YC: Young controls.

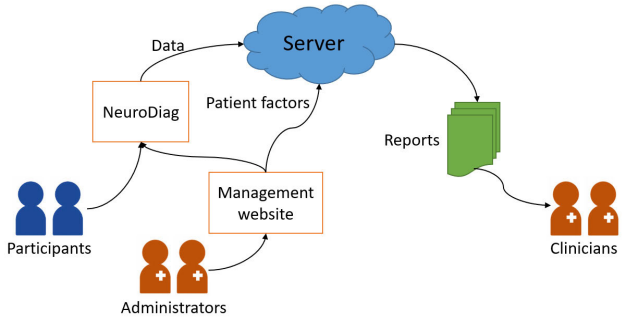


FIGURE 1. The overall block diagram of the proposed NeuroDiag for identifying people with PD using drawing and writing tasks.

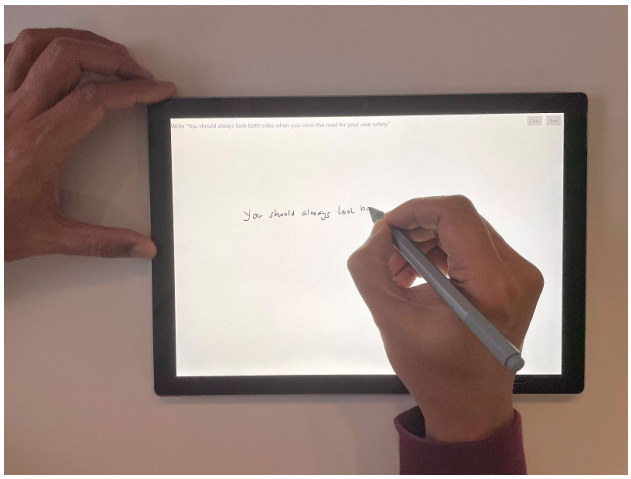


FIGURE 2. Photo of a participant using NeuroDiag doing Task 5 – reproducing a sentence.

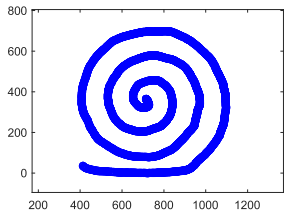
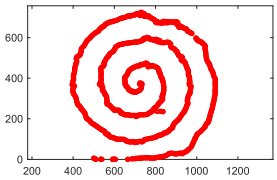
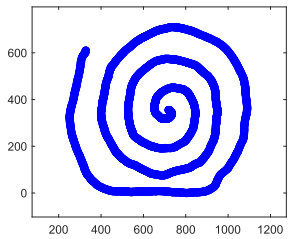
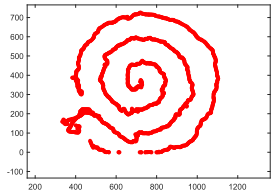
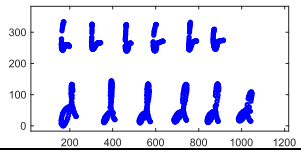
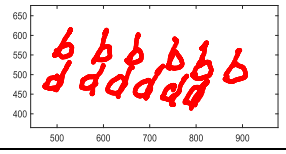
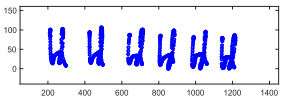
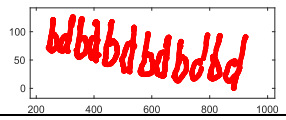
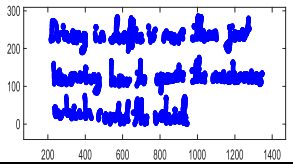
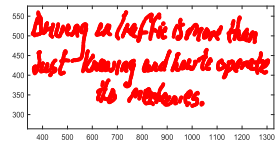
conducted on different days. In total, we got 39 handwriting records of age-matched controls (AMC-1: 26 records and AMC-2: 13 records). For the younger group (YC), 7 people participated, one of whom repeated the test on a different day and thus 8 records were collected. The participants' information is presented in Table 1.

B. NEURODIAG

NeuroDiag was developed at BioSignals Lab, School of Engineering, RMIT University. The software requires a patient ID number which is then attached to the registered email address. It is designed to direct the users to conduct a series of writing and drawing tasks and collect the movement data consisting of the x, y, the coordinates of contact between the screen and the stylus, the time stamp, and the pressure of the stylus. The data are sampled at 120 samples/second and after the completion of the tasks, this is transmitted to a server. The analysis of the data is done on the server. A report is generated automatically and in real-time which is transmitted to the user's registered email. The block diagram of the system is shown in Fig. 1. For the purpose of this test, a website was developed to record the patient factors and to generate patient IDs for use in NeuroDiag.

Data were collected using NeuroDiag when the participants performed six writing and drawing tasks on a Microsoft tablet with a stylus pen, as shown in Table 2 and Fig. 2. The first and second tasks consisted of drawing a spiral between dotted lines in a clockwise and an anti-clockwise direction, respectively. The drawing tasks are described with displayed dots that the user follows and thus the tasks are language-independent. The size and brightness of the dots were selected based on feedback from the users. The third task required repeated writing of the letter 'b' six times and 'd' six times which is considered a basic assessment of fine motor skills in writing [35]. The fourth task was writing 'bd' repeatedly as the writing strokes are distinctively affected by tremors [15]. Both tasks have letters with vertical strokes. Task 5 was

TABLE 2. Description and example of six handwriting tasks.

Task	Description	Sample Image Recorded from Age-Matched Healthy Control	Sample Image Recorded from PPD
Task 1	Sketching a dot-guided Archimedean spiral using the dominant hand		
Task 2	Sketching a dot-guided Archimedean spiral using the non-dominant hand		
Task 3	Writing the letters 'b' six times and 'd' six times.		
Task 4	Writing the letter 'bd' several times		
Task 5	Reproducing a sentence		
Task 6	Writing full name		

*The display of the full name was blurred to protect the participant's privacy.

the handwriting task that required attention and visuospatial memory compared with Tasks 3 and 4. Writing one's full name was Task 6, and this represents the natural writing style of the person.

C. DATA ANALYSIS

The software records the time stamp, the pressure of the stylus, and the location of the stylus on the tablet, which are sampled at 120 samples/second. Based on this data, the dynamics of pen pressure and handwriting kinematics, such as velocity, acceleration, and speed were extracted. The list of the extracted features from each task is presented in Table 3. Details of these parameters and analysis are in our previous study [15].

For statistical analysis, non-parametric Wilcoxon Rank Sum tests were applied to compare the group differences between people with PD and healthy people, and Spearman correlation analysis was conducted to investigate the association between handwriting patterns with the UPDRS-III data. Cohen's effect size was used to evaluate the differences in the handwriting patterns between PPD and AMC.

The classification model was based on statistical analysis and developed using data from PPD-GVH and AMC-1. The average and interquartile range of the most significant features extracted from each writing/sketching task from the two groups were used as the set of baseline parameters. In the testing phase, the extracted information was compared to the corresponding baseline. The classification was rule-based;

IEEE Journal of Translational Engineering in Health and Medicine

TABLE 3. Feature description of each writing/sketching task.

Task No	Feature name	Feature Description
Task 1	'F1',	Median value of pen pressure,
	'F2',	Median value of arctan,
	'F3',	Ratio of sample length and arctan,
	'F4',	Median of directional changes in X,
	'F5'	Maximum acceleration changes in Y.
Task 2	'F6',	Median value of pen pressure,
	'F7',	Median value of arctan,
	'F8',	Ratio of sample length and arctan,
	'F9',	Median of directional changes in X,
	'F10'	Maximum directional changes in Y.
Task 3	'F11',	Median value of pen pressure,
	'F12',	Average pen tip velocity in X,
	'F13',	Maximum pen tip acceleration in X,
	'F14',	Standard deviation of speed,
	'F15'	Average value of the speed.
Task 4	'F16',	Average pen tip velocity in Y,
	'F17',	Standard deviation of speed,
	'F18',	Average value of the speed,
	'F19',	Ratio of average velocity and max velocity,
	'F20'	Maximum pen tip acceleration in Y.
Task 5	'F21',	Median value of pen pressure,
	'F22',	Average pen tip velocity in Y,
	'F23',	Maximum pen tip acceleration in Y,
	'F24',	Standard deviation of speed,
	'F25'	Average value of the speed.
Task 6	'F26',	Median value of pen pressure,
	'F27',	Average pen tip velocity in Y,
	'F28',	Maximum pen tip acceleration in Y,
	'F29',	Standard deviation of speed,
	'F30'	Average value of the speed.

*X: horizontal direction; Y: vertical direction.

a minimum of 3 of the features should identify the sample as PD for it to be classified as ''PD''. The report contains the diagnosis outcome and also the confidence level to assist neurologists in their clinical decision-making.

III. RESULTS

PPD-GVH data and 26 records from the age-matched control group 1 data (AMC-1) were used to develop the model to differentiate between patients with PD and healthy people for NeuroDiag. The software was then tested in real-time on 38 participants (PPD-DN, AMC-2, and YC).

We compared handwriting patterns between PPD-GVH patients and the AMC-1 healthy group. The box plot of significant features (with P<0.01) is shown in Fig. 3. This figure shows that the kinematic features extracted from PD handwriting were smaller than age-matched healthy people. People with PD had a slower writing speed than healthy people in writing 'b' and 'd' (P=0.009), writing a sentence (P<0.001), and writing their full names (P<0.001). The figure reveals that the variability in the speed of PD patients was smaller than healthy people (P<0.001). In addition, the maximum pen tip acceleration in the vertical direction was also slower in PD patients compared to healthy people (P<0.001).

The effect size of the significant features is shown in Fig. 4. While the effect size for the average speed during writing 'b'

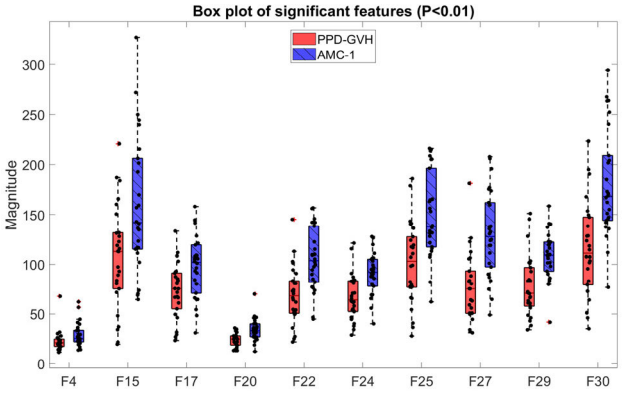


FIGURE 3. Box plot of significant features (P<0.01). F4: the median of directional changes in X during drawing spiral; F15: the average value of the speed during writing 'b' and 'd'; F17 and F20: the standard deviation of speed and the maximum pen tip acceleration in Y, respectively, during writing 'bd'; F22, F24, and F25: the average pen tip velocity in Y, the standard deviation of speed, and the average value of the speed, respectively, during writing a sentence; F27, F29, and F30: average pen tip velocity in Y, the standard deviation of speed, and the average value of the speed, respectively, during writing full name.

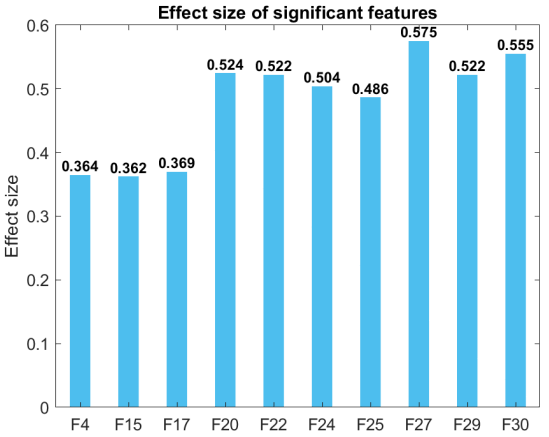


FIGURE 4. The effect size of the significant features. F4: the median of directional changes in X during drawing spiral; F15: the average value of the speed during writing 'b' and 'd'; F17 and F20: the standard deviation of speed and the maximum pen tip acceleration in Y, respectively, during writing 'bd'; F22, F24, and F25: the average pen tip velocity in Y, the standard deviation of speed, and the average value of the speed, respectively, during writing a sentence; F27, F29, and F30: average pen tip velocity in Y, the standard deviation of speed, and the average value of the speed, respectively, during writing full name.

and 'd' was the lowest (0.362), the average pen tip velocity in the Y (vertical) direction had the highest effect size (0.575).

The correlation between the significant features and UPDRS-III (PPD-GVH) is shown in Fig. 5. This figure shows that the highest correlation with UPDRS-III was that of the average speed during writing a sentence ($r = -0.42$).

With the baseline parameters, we achieved a sensitivity (True Positive) of 20/26 (76.92%) in correctly identifying the people with PD and specificity corresponding to a True Negative of 21/26 (80.77%) when validating the detecting model using PPD-GVH and AMC-1. When testing NeuroDiag in real-time using data from the PPD-DN, AMC-2, and

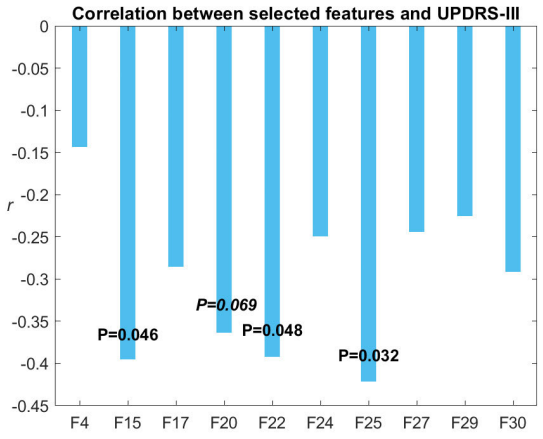


FIGURE 5. Spearman correlation between the selected features and UPDRS-III. F4: the median of directional changes in X during drawing spiral; F15: the average value of the speed during writing 'b' and 'd'; F17 and F20: the standard deviation of speed and the maximum pen tip acceleration in Y, respectively, during writing 'bd'; F22, F24, and F25: the average pen tip velocity in Y, the standard deviation of speed, and the average value of the speed, respectively, during writing a sentence; F27, F29, and F30: average pen tip velocity in Y, the standard deviation of speed, and the average value of the speed, respectively, during writing full name.

YC groups, the correct detection rate for PD (sensitivity) was 20/23 (86.96%) and the correct detection rate (specificity) for age-matched AMC-2 was 10/13 (76.92%), and for YC was 7/8 (87.50%).

IV. DISCUSSION

This paper introduces NeuroDiag, the software for the automated detection of PD using handwriting patterns. Data were collected from six drawing and writing tasks and uploaded to the server. Information about handwriting kinematics and pen pressure was extracted. The detection model was developed and tested using data from two different clinics. The results show that the software was able to detect PD with 86.96% sensitivity and 76.92% specificity when being tested in a real-time manner.

The software has the potential for telehealth. Data collected from the tablet were sent to the server where data analysis was conducted automatically. The software was designed to be used easily in general clinics. By registering as a user, a specialist can see the reports of patients and provide them with telehealth consultations.

Our findings in the current study are consistent with previous studies [14], [15], in which people with PD drew or wrote at a slower speed compared to healthy people. This is confirmed by the Spearman test, which showed a significant correlation between the average speed and average pen tip velocity of handwriting and the UPDRS-III data. The average speed during writing a sentence had a moderate correlation (absolute of $r > 0.40$) [36] with the rating scale. It is worth noticing that there were no significant differences between people with PD and healthy people when sketching a dot-guided Archimedean spiral using a non-dominant hand.

Fig. 3 shows that Task 5 and Task 6, writing a sentence and writing full name, respectively, are the most significant tasks for differentiating between people with PD and healthy people. Six out of 10 most significant features were extracted from these two tasks. This is consistent with the work by Drotár et al. [37], in which writing a sentence was proved to provide the highest correct detection rate of PD. As shown in Fig. 4, five features extracted from the two tasks have an effect size greater than the medium threshold [38]. These observations confirm the potential usefulness of handwriting for the detection of PD.

In this study, PD data from two different clinics were used to develop the software. For the age-matched control group, among eight people who participated in the test phase, six people were new to the study. Two people participated both in the developing phase and in the test phase, and one of them was misclassified as PD. Also in the real-time test phase, one healthy participant conducted the experiment 4 times, and NeuroDiag labeled them as PD one time. Although only data from age-match healthy people were used to develop the model, we achieved a good correction rate of 7/8 (87.50%) when testing the software with young people. This demonstrates the robustness of NeuroDiag and its potential for assisting neurologists and clinicians with fast and accurate screening of people with Parkinson's disease.

There are three major limitations in the current work. Firstly, the baseline parameters were built from a small number of participants. However, NeuroDiag has shown its potential with a good correct detection rate when testing on a different data set in real-time. Nevertheless, a large number of participants is required to validate the robustness of NeuroDiag. The classification model could be then updated with shadow machine learning models or with deep learning models. Secondly, the software was tested using only one tablet model, Microsoft Surface Pro. While other tablet models have the necessary parameters, however, this has not been tested. Thirdly, the security of the system was not tested.

This study was a cross-sectional study, where the clinical data and the NeuroDiag data were recorded during the same patient visit. Future research would be related to testing the feasibility of NeuroDiag to estimate the effect of medication and the severity of the disease. A longitudinal study could lead to the use of NeuroDiag to monitor the progression of the disease. Including the meta-data of the user may also improve the model performance. There is also the need for developing an explainable artificial intelligence model, where the diagnosis outcome is interpretable with support from the baseline parameters and patient-related metadata.

V. CONCLUSION

NeuroDiag was developed and tested using multi-centre data. From six handwriting tasks, the software was able to differentiate between people with PD, most of whom were in the early stage of the disease, and healthy people with a sensitivity of 86.96% and a specificity of 76.92%. These classification results show that NeuroDiag could be used to

assist neurologists in assessing their patients. The proposed server-based model shows the feasibility of NeuroDiag being used for telehealth. These show that NeuroDiag has the potential for translation into clinical practice, which could promote the development of a community-wide screening program for Parkinson's disease, especially in remote areas.

DATA AVAILABILITY

The de-identified data will be made available on request to the corresponding author.

REFERENCES

[1] Y. Xu et al., "Neurotransmitter receptors and cognitive dysfunction in Alzheimer's disease and Parkinson's disease," *Prog. Neurobiol.*, vol. 97, no. 1, pp. 1–13, Apr. 2012.

[2] R. B. Postuma and J. Montplaisir, "Predicting Parkinson's disease—Why, when, and how?" *Parkinsonism Rel. Disorders*, vol. 15, pp. S105–S109, Dec. 2009.

[3] S. Sveinbjornsdottir, "The clinical symptoms of Parkinson's disease," *J. Neurochem.*, vol. 139, pp. 318–324, Oct. 2016.

[4] C. G. Goetz et al., "Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement Disorders*, vol. 23, no. 15, pp. 2129–2170, Nov. 2008.

[5] Q. C. Ngo, M. A. Motin, N. D. Pah, P. Drotár, P. Kempster, and D. Kumar, "Computerized analysis of speech and voice for Parkinson's disease: A systematic review," *Comput. Methods Programs Biomed.*, vol. 226, Nov. 2022, Art. no. 107133.

[6] N. Miller, E. Noble, D. Jones, and D. Burn, "Life with communication changes in Parkinson's disease," *Age Ageing*, vol. 35, no. 3, pp. 235–239, 2006.

[7] C. A. Haaxma et al., "Gender differences in Parkinson's disease," *J. Neurol., Neurosurgery, Psychiatry*, vol. 78, no. 8, pp. 819–824, 2007.

[8] M. Thomas, A. Lenka, and P. Kumar Pal, "Handwriting analysis in Parkinson's disease: Current status and future directions," *Movement Disorders Clin. Pract.*, vol. 4, no. 6, pp. 806–818, Nov. 2017.

[9] J. Jankovic, "Parkinson's disease: Clinical features and diagnosis," *J. Neurol., Neurosurg. Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.

[10] A. R. E. Potgieser, E. Roosma, M. Beudel, and B. M. de Jong, "The effect of visual feedback on writing size in Parkinson's disease," *Parkinson's Disease*, vol. 2015, pp. 1–4, Jun. 2015.

[11] M. E. Morris, J. McGinley, F. Huxham, J. Collier, and R. Iansek, "Constraints on the kinetic, kinematic and spatiotemporal parameters of gait in Parkinson's disease," *Hum. Movement Sci.*, vol. 18, nos. 2–3, pp. 461–483, Jun. 1999.

[12] P. Drotár, 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," *Artif. Intell. Med.*, vol. 67, pp. 39–46, Feb. 2016.

[13] A. Letanneux, J. Danna, J.-L. Velay, F. Viallet, and S. Pinto, "From micrographia to Parkinson's disease dysgraphia," *Movement Disorders*, vol. 29, no. 12, pp. 1467–1475, Oct. 2014.

[14] P. Zham, D. K. Kumar, P. Dabnichki, S. P. Arjunan, and S. Raghav, "Distinguishing different stages of Parkinson's disease using composite index of speed and pen-pressure of sketching a spiral," *Frontiers Neurol.*, vol. 8, p. 435, Sep. 2017.

[15] P. Zham, S. P. Arjunan, S. Raghav, and D. K. Kumar, "Efficacy of guided spiral drawing in the classification of Parkinson's disease," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 5, pp. 1648–1652, Sep. 2018.

[16] K. Sarin et al., "A three-stage fuzzy classifier method for Parkinson's disease diagnosis using dynamic handwriting analysis," *Decis. Anal. J.*, vol. 8, Sep. 2023, Art. no. 100274.

[17] K. Kumar and R. Ghosh, "Parkinson's disease diagnosis using recurrent neural network based deep learning model by analyzing online handwriting," *Multimedia Tools Appl.*, vol. 83, no. 4, pp. 11687–11715, Jan. 2024.

[18] P. Varalakshmi, B. T. Priya, B. A. Rithiga, R. Bhuvaneaswari, and R. Sakthi Jaya Sundar, "Diagnosis of Parkinson's disease from hand drawing utilizing hybrid models," *Parkinsonism Rel. Disorders*, vol. 105, pp. 24–31, Dec. 2022.

[19] R. Gupta, S. Kumari, A. Senapati, R. K. Ambasta, and P. Kumar, "New era of artificial intelligence and machine learning-based detection, diagnosis, and therapeutics in Parkinson's disease," *Ageing Res. Rev.*, vol. 90, Sep. 2023, Art. no. 102013.

[20] J. C. Vásquez-Correa, T. Arias-Vergara, J. R. Orozco-Arroyave, B. Eskofier, J. Klucken, and E. Nöth, "Multimodal assessment of Parkinson's disease: A deep learning approach," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 4, pp. 1618–1630, Jul. 2019.

[21] S. M. Abdullah et al., "Deep transfer learning based Parkinson's disease detection using optimized feature selection," *IEEE Access*, vol. 11, pp. 3511–3524, 2023.

[22] Z. Galaz et al., "Comparison of CNN-learned vs. Handcrafted features for detection of Parkinson's disease dysgraphia in a multilingual dataset," *Frontiers Neuroinform.*, vol. 16, May 2022, Art. no. 877139.

[23] W. J. Gordon and A. D. Stern, "Challenges and opportunities in software-driven medical devices," *Nature Biomed. Eng.*, vol. 3, no. 7, pp. 493–497, Jul. 2019.

[24] S. Gerke, B. Babic, T. Evgeniou, and I. G. Cohen, "The need for a system view to regulate artificial intelligence/machine learning-based software as medical device," *NPJ Digit. Med.*, vol. 3, no. 1, p. 53, Apr. 2020.

[25] K. N. Vokinger, S. Feuerriegel, and A. S. Kesselheim, "Mitigating bias in machine learning for medicine," *Commun. Med.*, vol. 1, no. 1, p. 25, Aug. 2021.

[26] M. Linares-del Rey, L. Vela-Desojo, and R. Cano-de la Cuerda, "Mobile phone applications in Parkinson's disease: A systematic review," *Neurología English Ed.*, vol. 34, no. 1, pp. 38–54, Jan. 2019.

[27] L. Fraiwan, R. Khnouf, and A. R. Mashagbeh, "Parkinson's disease hand tremor detection system for mobile application," *J. Med. Eng. Technol.*, vol. 40, no. 3, pp. 34–127, 2016.

[28] N. Kostikis, D. Hristu-Varsakelis, M. Arnaoutoglou, and C. Kotsavasiloglou, "Smartphone-based evaluation of parkinsonian hand tremor: Quantitative measurements vs clinical assessment scores," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 906–909.

[29] B. M. Bot et al., "The mPower study, Parkinson disease mobile data collected using ResearchKit," *Sci. Data*, vol. 3, no. 1, Mar. 2016, Art. no. 160011.

[30] H. Kim et al., "Unconstrained detection of freezing of gait in Parkinson's disease patients using smartphone," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 3751–3754.

[31] B. P. Printy et al., "Smartphone application for classification of motor impairment severity in Parkinson's disease," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 2686–2689.

[32] S. Arora et al., "Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study," *Parkinsonism Rel. Disorders*, vol. 21, no. 6, pp. 650–653, Jun. 2015.

[33] S. Aghanavesi, M. Memedi, M. Dougherty, D. Nyholm, and J. Westin, "Verification of a method for measuring Parkinson's disease related temporal irregularity in spiral drawings," *Sensors*, vol. 17, no. 10, p. 2341, Oct. 2017.

[34] R. B. Postuma et al., "MDS clinical diagnostic criteria for Parkinson's disease," *Movement Disorders*, vol. 30, no. 12, pp. 1591–1601, Oct. 2015.

[35] W. G. K. Cobbah and M. C. Fairhurst, "Computer analysis of handwriting dynamics during dopamimetic tests in Parkinson's disease," in *Proc. 26th Euromicro Conf., EUROMICRO Inform., Inventing Future*, Sep. 2000, pp. 414–418.

[36] P. Schober, C. Boer, and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesthesia Analgesia*, vol. 126, no. 5, pp. 1763–1768, 2018.

[37] P. Drotár, J. Mekyska, Z. Smékal, I. Rektorová, L. Masarová, and M. Faundez-Zanuy, "Prediction potential of different handwriting tasks for diagnosis of Parkinson's," in *Proc. E-Health Bioeng. Conf. (EHB)*, Nov. 2013, pp. 1–4.

[38] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. New York, NY, USA: Routledge, 1988.