What Can Gestures Tell? Detecting Motor Impairment in Early Parkinson's from Common Touch Gestural Interactions

Feng Tian[1,2], Xiangmin Fan[1,2†], Junjun Fan[1,2], Yicheng Zhu[3], Jing Gao[1,2], Dakuo Wang[4], Xiaojun Bi[5], Hongan Wang[1,2]

[1] State Key Laboratory of Computer Science and Beijing Key Lab of Human-Computer Interaction, Institute of Software, Chinese Academy of Sciences, Beijing, China

[2]School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing, China

[3]Peking Union Medical College Hospital, Beijing, China

[4]IBM Research, Yorktown Heights, New York, United States

[5]Department of Computer Science, Stony Brook University, Stony Brook, New York, United States

{tianfeng, xiangmin, junjun2015, hongan}@iscas.ac.cn, zhuyc@pumch.cn, gaojing162@mails.ucas.ac.cn, dakuo.wang@ibm.com, xiaojun@cs.stonybrook.edu

†: Corresponding author

ABSTRACT

Parkinson's disease (PD) is a chronic neurological disorder causing progressive disability that severely affects patients' quality of life. Although early interventions can provide significant benefits, PD diagnosis is often delayed due to both the mildness of early signs and the high requirements imposed by traditional screening and diagnosis methods. In this paper, we explore the feasibility and accuracy of detecting motor impairment in early PD via sensing and analyzing users' common touch gestural interactions on smartphones. We investigate four types of common gestures, including flick, drag, pinch, and handwriting gestures, and propose a set of features to capture PD motor signs. Through a 102-subject (35 early PD subjects and 67 age-matched controls) study, our approach achieved an AUC of 0.95 and 0.89/0.88 sensitivity/specificity in discriminating early PD subjects from healthy controls. Our work constitutes an important step towards unobtrusive, implicit, and convenient early PD detection from routine smartphone interactions.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; *Empirical studies in ubiquitous and mobile computing*; Ubiquitous and mobile computing design and evaluation methods.

KEYWORDS

Parkinson's disease (PD), touch gestures, passive monitoring.

ACM Reference Format:

Feng Tian[1,2], Xiangmin Fan[1,2†], Junjun Fan[1,2], Yicheng Zhu[3], Jing Gao[1,2], Dakuo Wang[4], Xiaojun Bi[5],, Hongan Wang[1,2]. 2019. What Can Gestures Tell? Detecting Motor Impairment in Early Parkinson's from Common Touch Gestural Interactions. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3290605.3300313

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI 2019, May 4–9, 2019, Glasgow, Scotland Uk

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-5970-2/19/05…$15.00

https://doi.org/10.1145/3290605.3300313

1 INTRODUCTION

Parkinson's disease (PD) is the most common movement disorder and the second most common neurodegenerative disease [56]. Age-standardized incidence rates of PD in population-based studies from Europe and the USA range from 8.6 to 19 per 100,000 inhabitants [11]. The prevalence of PD among elderly is even much higher (e.g., 9.5 per 1,000 people 65 or older [22]). PD is characterized by both motor symptoms (e.g., tremor, rigidity, bradykinesia, and postural instability) and non-motor symptoms (e.g., cognitive alteration and sleep disturbances) [57], which severely affects patients' well-being and quality of life. Besides, PD brings economic burdens to both patients and the society. The total annual cost of the disease was estimated to be 13,800 Euros per individual in a Swedish survey [19], and the costs increase significantly with disease severity.

While early treatment can produce significant benefits for patients, the clinical diagnosis of PD is usually delayed in practice. Due to the mildness of many early signs, patients may not undergo clinical examinations during early stages.

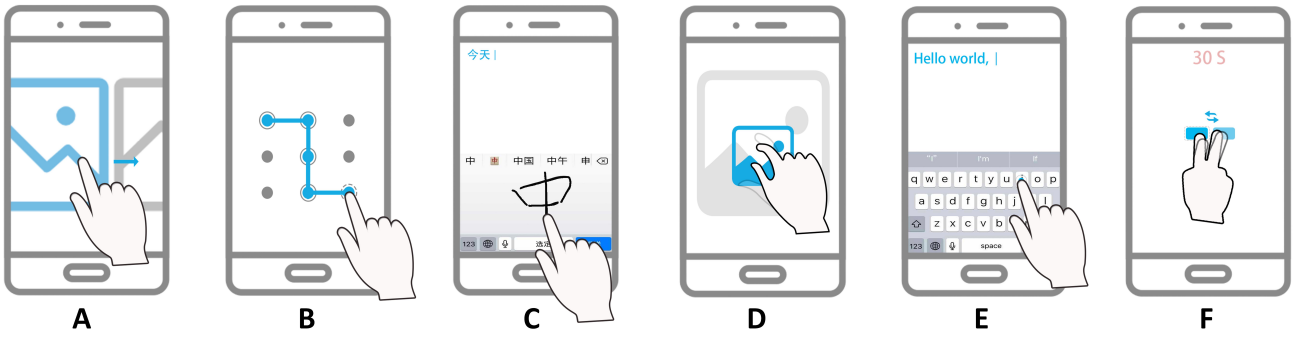

Figure 1: Common touch gestures explored in this study. A: flick gestures; B: drag gestures; C: handwriting gestures; D: pinch gestures; E: tap gestures (typing); F: alternating finger tapping (AFT, included as a clinical reference).

Moreover, the traditional procedure of PD screening and diagnosis rely heavily on the clinical judgement of neurologists. The Unified Parkinson's Disease Rating Scale (UPDRS) [17] is the most widely used scale to evaluate subjects' overall conditions. In particular, motor status is assessed in the clinician-scored motor section (UPDRS Part III), based on the evaluation of the observed performance in a series of standardized motor tasks. Such evaluation suffers from both inter-rater variability [57] and recall-bias of participants [42]. Besides, the need for a specialist and the presence of the subject at the clinic limit the frequency of screening and monitoring of PD symptoms. As a result, the disease may be undiagnosed for many years. In a prior study [54], researchers found that most presentations of motor impairments, such as tremor and bradykinesia, occurred within 2 years before the first clinical diagnosis of PD. They also showed that the incidence of tremor was already higher in PD subjects compared with the control group at up to 10 years before diagnosis.

Researchers, clinicians, as well as patients share a keen interest in the development of more accessible tools that can provide objective PD assessment, which can lead to timely diagnosis and consequent improvement of the prospective patient's life quality via early therapeutic interventions [40]. A variety of techniques have been proposed to achieve this goal. One of the main trends is the application of IMUs (Inertial Measurement Units) mounted on human bodies for motor assessment [1, 2, 8, 12, 27, 30, 43, 44, 49]. During recent years, the use of commodity devices such as smartphones [1, 2, 4, 24] has been more preferred over systems using specialized sensors, since such non-obtrusive approaches show more potential for transferring PD screening and monitoring in the daily life. The mPower study [6] is one of the most well-known studies of longitudinal and large-scale smartphone-based data collection for PD research. Participants (PD patients and healthy subjects) perform certain tasks, such as touchscreen tapping, memory, voice, and gait (walk) tests, to contribute data remotely. Although it has involved more than 9,000 participants, the drop-out rate was high. Bot et al. [6] reported that there were less than 10% participants who performed the tapping test for more than 5 days. This reveals a common limitation of prior work-they require users' explicit and active participation, in the sense that users need to be reminded to mount/wear sensing devices and perform certain tests. Such requirements lead to reduced compliance due to the lack of sustained motivation, especially when the symptoms are not evident or even not noticeable.

We propose a solution that takes advantage of the pervasiveness and ubiquity of touchscreens in the smartphone era. Performing touch gestures is a daily interaction task of modern humans, yet making precise finger movement on touchscreens requires users' fine motor skills as well as motor coordination abilities. Since PD motor phenotype is described by tremor, slowness, and rigidity, which may affect the unconstrained finger performance while interacting with smartphones, we hypothesize that we can build a machine learning model which can discriminate PD patients from healthy subjects via sensing and analyzing their touch behaviors. Compared with previous approaches that require users to mount sensing equipment, explicitly launch monitoring apps, and spend an uninterrupted amount of time in data collection, our approach has the potential to make the detection more "*implicit*" since it can be a side effect of performing everyday interaction tasks.

In this paper, we explore the feasibility and accuracy of detecting motor impairment in early PD through common touch gestural interactions. Specifically, we investigate four types of common gestures including flick (quickly brush surface with fingertip), drag (move fingertip over surface without losing contact), pinch (touch surface with two fingers and move them together/apart), and handwriting gestures

(i.e. multi-stroke free-form path gestures). We propose a set of touch gesture features including path/trajectory-based features, time-based features, pressure-based features and IMU-based features, to capture PD motor signs. We also conducted a thorough exploration of the feature space to gain a holistic understanding about the discriminative power of each category of features. Through a 102-subject study (35 early PD patients and 67 age-and-gender-matched controls), our approach achieved an AUC (area under the ROC curve) of 0.95 and 0.89/0.88 sensitivity/specificity in discriminating PD and control subjects. In comparison, the results of the alternating finger tapping (AFT) test, a well-established motor test in clinics, were 0.83 AUC and 0.86/0.69 sensitivity/specificity. These results shed light on the feasibility of unobtrusive and implicit early PD detection from everyday smartphone interactions.

The contributions of our research are three fold. First, we propose a new methodology to quantify PD motor impairment through common touch gestural interactions. To the best of our knowledge, this is the first exploration of using touchscreen *gesturing* as a cue for early PD detection. Second, we conduct a thorough investigation of both common touch gestures and their corresponding feature spaces, and provide a set of fundamental findings which can provide a solid base to rely on for future researchers and practitioners. Third, we contribute a touch gesture dataset consisting of data collected from 35 early-PD and 67 control subjects, which can serve as a common ground for future research and performance comparison of different approaches.

2 RELATED WORK

Emerging Technologies for Objective and Quantitative PD Assessment

As we discussed earlier, traditional measurement methods of PD, e.g., UPDRS [17], rely heavily on raters' experience and subjects' self-reports, thus suffer from inter-rater variability and subjects' recall bias. The need for objective, quantitative, and more accessible assessment has resulted in an exponential development of technologies in this field [1, 2, 4, 8, 12, 14, 16, 24, 27, 29, 30, 43, 44, 49, 52, 53]. Since motor impairment is the most dominant symptom of PD, the overwhelming majority of works focus on monitoring subjects' motor functions. Ghika et al. [16] were among the earliest researchers to explore the use of sensing technologies (e.g., accelerometers mounted on the hand) to quantify and monitor motor abnormalities including tremor, bradykinesia, and rigidity. The advances in sensor miniaturization and wireless technology allows recent researchers to develop solutions based on unmodified commodity devices such as smartphones [1, 2, 4, 24] and Google Glass [36, 37]. In addition to accelerometers and gyroscopes (alone or combined, which we categorize as inertial measurement units, IMUs) which were the primary choice of sensors of existing approaches, recent works also explored using Kinect sensors [14], microphones [2], and ad hoc solutions for objective PD assessment. The intended applications range from the improvement of diagnosis to the assessment of PD progression and the evaluation of therapy efficacy [40]. Please refer to [53] for a systematic review of 848 articles published from 2005 to 2015, which provides a panoramic overview of technological innovations in PD evaluation of the past decade.

Smartphone-based Assessment Methods

We pay special attention to smartphone-based assessment methods [1–3, 8, 12, 24, 27–30, 34, 43–45, 48, 49], which are more related to our work. While most of them focused on detecting motor abnormalities via the embedded IMUs [1, 2, 8, 12, 27, 30, 43, 44, 49], a few studies leveraged touchscreen as a source of data to quantify PD signs [3, 24, 29, 49]. For example, Lee et al. [29] implemented and validated finger-tapping (FT) tests [41] on smartphones. FT tests are standardized finger-movement tests to detect and quantify psychomotor dysfunction and were traditionally performed on mechanical tappers. When transferred to touchscreen tappers, they achieved a 0.92 AUC when distinguishing PD patients and controls. Similarly, the mPower study [6] included alternating finger-tapping (AFT) tests [23] on smartphones. AFT is one of the varieties of FT, in which the tested subject has to alternatively press two specific buttons as fast as possible during a predefined time. The test is repeated for both hands and the final score is the average number of pressed keys between the two. In addition to AFT tests, the mPower study also included memory tests, audio tests, and gait (walk) tests. For instance, in the audio test, a participant is asked to "say 'Aaaaah' into the microphone for as long as you can". The audio will be collected for further analysis. Similar with prior test-based approaches, the mPower study also suffers from high drop-out rates (e.g., >90% subjects dropped out AFT tests within 5 days) since the data collection requires users' active participation and it is therefore, subject to adherence. Moreover, according to the *observer effect* (aka Hawthorne effect [39]) which indicates that individuals modify some aspect of their behavior in response to their awareness of being observed, data collected in such tests can hardly reflect the natural behavior of the user, which may further impact the results to certain extent.

To address these problems, recent research has explored using keystroke dynamics (i.e. timing information associated with keystrokes) of typing activities to classify subjects as having PD or not [3, 4, 24]. Arroyo-Gallego et al. [4] used statistics of hold time (time between pressing and releasing a key), flight time (delay between consecutive key presses), and

other raw typing metrics generated when subjects typed on a mechanical keyboard and achieved a 0.83 AUC in discriminating early PD subjects and controls. They further validated this approach using a touchscreen smartphone and achieved a 0.91 AUC with the best performing feature [3]. Iakovakis et al. [24] further extended this approach by including more features (e.g., touch pressure) and achieved a 0.92 AUC (the method proposed in [3] achieved a 0.82 AUC on the same dataset).

Although highly related, we go beyond existing literature by presenting the first exploration of detecting early PD signs from common touch gestural interactions. Whereas typing behaviors can reflect one's fine motor control abilities in the form of *repeated target selection*, gesturing behaviors can reflect one's abilities in making *precise finger movements*. By including one recent work [24] as a baseline method in our evaluation, we demonstrate through empirical evidence that the combination of typing and gesturing can outperform either of them solely. Moreover, introducing gesturing as another cue for PD detection can provide a better coverage regarding usage scenarios. Besides, we explore an enriched feature set covering not only time-based features and pressure-based features, but also path-based features and IMU-based features, which have not been fully covered by prior works. We believe such exploration can help the community gain a holistic understanding of the effects caused by PD motor symptoms on touch gesturing.

Utilizing Side Effects of Touch Interactions

Recent studies in psychology literature have shown that touch behaviors in a social context (person-person communication) can convey not only the valence of an emotion but also the type of that emotion (e.g., happy or upset) [21]. Motivated by such psychological findings, researchers in the HCI community have explored using the touch modality to capture users' emotional states [5, 15, 25, 31]. Bailenson et al. [5] found that emotions can be communicated through a two degrees of freedom force feedback joystick. Khanna & Sasikumar [25] found that keyboard typing behavior can be affected by users' emotional states and developed a system that can discriminate positive, neutral and negative states based on keystroke behaviors with up to 87.7% accuracy. Lv et al. [31] achieved 93.4% accuracy in discriminating six emotions (neutral, anger, fear, happiness, sadness, and surprise) via a keyboard with pressure sensors. Gao et al. [15] explored detecting players' emotional states in a naturalistic setting of touchscreen-based games via analyzing their finger stroke behaviors, and achieved up to 77% accuracy in discriminating between four emotional states.

In addition to emotional states, researchers have also explored detecting *situational impairments* imposed by inebriation via touch-based interaction tasks [33], which is more in line with our work. In that work, researchers proposed five drunk user interfaces (DUIs), including typing, swiping, balancing & heart rate, simple reaction, and choice reaction. Among these tasks, the typing and swiping interfaces are particularly related to our research. As we illustrate in later sections, this work has partially informed the selection of features in our study; nevertheless, there are several major distinctions between these two works. First, DUIs focus on detecting motor abnormalities caused by inebriation while our work focuses on detecting PD motor signs, which may differ from each other. Second, in addition to typing and swiping (i.e. tracing pattern passwords), we explore a wider range of gestures such as flick, pinch, and handwriting gestures. Third, DUIs were designed to be dedicated task-based interfaces while we intend to collect and analyze users' data when performing common touch gestures in a task-free manner.

3 EXPERIMENT

In the modern era of smartphones, touch gestural interactions are common tasks that users perform every day. Performing touch gestures involve precise, coordinated and successive finger and hand movements. While prior clinical experiments showed that PD motor phenotype described by bradykinesia/slowness, lack of spontaneous movement, rigidity and tremor can reduce the coordination of fingers in fine motor control [35, 58], we believe such clinical picture can also affect the finger performance while interacting with smartphone touchscreens. Therefore, analyzing the quantitative information arising from this type of interaction is the focus of our research. Moreover, such information can be captured unobtrusively and implicitly in the background during routine interaction scenarios. In this way, it can better reflect the natural behavior of the user while reducing the *observer effect* [39].

In this study, we investigate the possibility of detecting motor impairments in early PD via analyzing users' touch gestural behaviors. Specifically, we introduce a set of numerical features derived from common touch gestures and investigate the characteristic PD patterns that can facilitate the detection and quantification of the motor signs related to this disease. The data in this study was collected from 102 subjects (35 early PD patients and 67 healthy controls). In the rest of this section, we describe the selection of gestures and the extraction of features, and cite a subset of clinical experiments and prior studies that informed them. We also describe the participants and the data collection procedure of this study.

Gestures

We explore four types of common touch gestures including flick, drag, handwriting, and pinch (Figure 1). These gestures

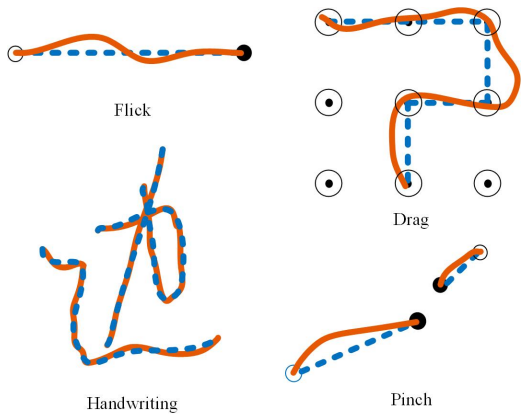

Figure 2: Ideal finger paths (dashed lines, hidden from the user) and actual finger paths (solid paths).

combined can provide a good coverage of daily touch gestural interactions. Previous studies on PD research suggests that PD motor impairments can affect patients' fine-motor coordination and control abilities [35, 58]; to the best of our knowledge, though, there has been no work that has quantitatively analyzed the effect of PD on gestural interactions performed on touchscreens. In addition, we also include two external references in the evaluation: typing [24] and the AFT test (a well-established clinical test [23]).

Flick. A flick is a unidirectional touch gesture that requires the user to quickly brush the screen surface with her fingertip. It is typically characterized by high *speed* and high degree of *straightness*. We hypothesize that the bradykinesia/slowness, tremor, and other PD symptoms can affect patients' efficiency when performing flick gestures regarding both finger movement speed and path efficiency (e.g., straight or zigzag path).

Drag. A drag gesture requires the user to move fingertip over touchscreen surface without losing contact. Typically, each trial has both a desired starting position and a desired end position, thus it can be considered as a steering task with explicit or implicit tunnels. When comparing the actual finger trajectory with the ideal path (e.g., the shortest path between the starting and end positions), one hypothesis is that PD patients would move their fingers less efficiently than healthy subjects. To our knowledge, the effect of PD motor impairment on drag gestures performed on touchscreens has yet to be explored, but there have been related studies involving the effect of inebriation. For example, Mariakakis et al. [33] demonstrated that motor abnormalities caused by inebriation can affect both finger speed and trajectory efficiency when tracing pattern passwords on an Android 3x3 lock screen.

Handwriting. Handwriting gestures can be considered as free-form multi-stroke path gestures. Handwriting input is widely adopted by smartphone users worldwide. For example, Google Handwriting Input[1] alone has attracted more than 10 million installs to date. Compared with drag gestures, handwriting gestures can reflect less-constrained finger movements. Prior studies showed that PD symptoms could affect keystroke temporal dynamics and touch pressures when typing with an onscreen Qwerty keyboard [24], we explore whether and how PD symptoms affect handwriting behaviors in terms of path smoothness, movement speed, and touch pressure.

Pinch. A pinch gesture requires the user to touch screen surface with two fingers and move them closer or apart, which involves multi-finger coordination. While the well-known AFT test [23] measures one's ability in performing alternating and successive finger tapping tasks, pinch behaviors can reflect one's ability in making simultaneous, continuous, and precise finger movements. Thus we believe certain characteristic PD patterns can be detected via sensing and analyzing users' pinch behaviors.

Tap. Prior research has investigated PD patients' tap behaviors in both AFT tests [29] and typing activities [3, 24]. We include both in this study as external references and baseline methods. In the AFT test, a user alternately taps two separate buttons on the screen by using her index finger and middle finger on the same hand for 30 seconds. The user needs to repeat the test for both hands and the final score is computed as the average number of buttons pressed between two hands. Despite its simplicity, AFT has been widely used to quantify upper limbs dexterity for PD evaluation [41]. Therefore, it can also serve as a clinical reference in the study. In addition to the AFT test, we also include onscreen Qwerty typing-based approach as another external reference, followed the implementation of a recently published work by Iakovakis et al. [24].

Features

We explore four categories of features and investigate whether and to what extent they can capture the characteristic patterns caused by PD motor symptoms. Table 1 shows a complete listing of each feature. We cite the related literature that informed the calculations of certain features in this study, and describe how these features are adapted for use in our scenario. Although some of the features were originally proposed by prior researchers in other tasks, we focus on analyzing their discriminative power in differentiating early PD and normal subjects in this study.

[1]https://play.google.com/store/apps/details?id=com.google.android.apps.handwriting.ime&hl=en_US

Table 1: Features explored in this study.

Features	Descriptions
Path-based Features	
Ideal Path Crossing	The number of times that the actual path crosses the ideal path [32].
Path Direction Change	The number of times that the finger changes direction relatively to the ideal path [32].
Path Deviation Variability	The std of the distances between points on actual path and ideal path [32].
Path Error	The mean absolute deviation of the points on the actual path from the ideal path [32].
Path Offset	The mean deviation of the points on the actual path from the ideal path [32].
Curvature	The mean directional change in degrees of each point on the actual path [38].
Time-based Features	
Speed (min/max/mean/median/std)	The min/max/mean/median/std of swiping speed in pixels/s [33].
Acceleration (min/max/mean/median/std)	The min/max/mean/median/std of swiping acceleration in pixels/s^2 [33].
Speed Jerk (min/max/mean/median/std)	The min/max/mean/median/std of swiping jerk in pixels/s^3 [33].
Pressure-based Features	
Pressure (min/max/mean/median/std)	The min/max/mean/median/std of pressure values [15].
Pressure Change (min/max/mean/median/std)	The min/max/mean/median/std of first derivative of pressure values.
Pressure Jerk (min/max/mean/median/std)	The min/max/mean/median/std of second derivative of pressure values.
IMU-based Features	
Amplitude	The root mean square of all the input-squared values [47].
Voluntary Movements Power	The relative power of voluntary movements frequencies (1-3.5Hz) [20].
Pathological Tremor Power	The relative power of pathological tremor frequencies (3.5-7.5Hz) [20].
Physiological Tremor Power	The relative power of physiological tremor frequencies (7.5-15Hz) [20].
Mean Frequency	The mean frequency of the power spectrum [47].
Median Frequency	The particular frequency that divides the spectrum into two parts of equal area [47].
Additional Features for Handwriting Gestures	
Finger Stroke Interval	The mean interval between two consecutive finger strokes.
Interval Variability	The std of intervals between two consecutive finger strokes.
Additional Features for Pinch Gestures	
Speed Ratio	The mean speed ratio of the two moving fingers.
Pressure Ratio	The mean pressure ratio of the two moving fingers.

Path-based features. Here we hypothesize that motor impairments in PD can affect the efficiency in term of finger trajectories (e.g., zigzag path vs. smooth path). To measure the path efficiency, we need to know the ideal path first. For flick, drag and pinch gestures, the most efficient path is the straight lines connecting the starting point, the end point, and the crossing goals in between (if existed). For handwriting paths, we applied a mean filter with an averaging window length of 5 to generate a smoothed path as the ideal path [51]. Figure 2 shows the ideal paths and an user's actual paths. Then we applied the path accuracy features introduced by Mackenzie et al. [32], which evaluates how a trajectory deviates from the ideal path between them. For instance, "path error" measures the absolute orthogonal distance between user's trajectory and the most efficient path. Since the stroke lengths vary even when the same user performs the same gesture twice, we normalize the feature values based on stroke length. In addition to path accuracy features, we also included path smoothness metrics (e.g., curvature), inspired by prior research which demonstrated that PD symptoms can affect the smoothness of patients' eye and head movements [59].

Time-based features. Since bradykinesia (i.e. slowness) is one of the dominant manifestations of PD, time-based information has been widely adopted by prior research for PD quantification. For example, Arroyo-Gallego et al. [3]

discriminated PD and control subjects via analyzing their typing behaviors based on timing features such as holding time and flight time. In addition to PD assessment, Mariakakis et al. [33] used finger velocity, acceleration, and other time-based measurements in estimating the blood alcohol level. Informed by prior research, we include a set of time-based features, including finger movement speed, acceleration, jerk, as well as the summary statistics such as their min, max, mean, median, and standard deviation values.

Pressure-based features. Touch pressure-based features were found to be informative by prior studies for estimating situational motor impairments introduced by inebriation [33], and for discriminating emotions [15]. The most related finding is presented by Iakovakis et al. [24], which showed that pressure features can facilitate the discrimination of PD and control subjects during typing activities. Informed by these findings, we calculate a set of features based on the pressure sequence captured by the touchscreen, such as the summary statistics of the normalized pressure values (i.e. 0-1), and statistics that can describe the smoothness of pressure change in temporal domain.

IMU-based features. In addition to touchscreen, we also include the embedded IMU as another source to collect data. Relying on IMU for PD quantification has been a long tradition of PD research [1, 2, 16, 20, 49]. However, as far as we know, there is no work on analyzing IMU data collected during touch gestural interactions. We include both generic features (e.g., amplitude) and features that can reflect tremor symptoms (e.g., power of tremor frequencies).

We calculate all four categories of features discussed above for each type of gestures. In addition, we also include some unique features for handwriting and pinch gestures. For handwriting, we include features that can reflect the finger stroke dynamics (e.g., mean/std of the time interval between two consecutive finger strokes), informed by previous clinical findings that PD patients lack on rhythm stability of their finger movements [26, 35]. For pinch gestures, we include features that can reflect the coordination of the two moving fingers, such as the ratio of speeds as well as the ratio of pressures of the two fingers.

Participants

We recruited two groups of participants. The early PD subjects group consisted of 35 subjects. All of them were at early stages (Hoehn-Yahr stages I or II[2], mean UPDRS Part III score/std 8.4/3.7) and recently diagnosed (mean disease onset years/std 2.3/1.9). The diagnosis was confirmed by at least two expert physicians before the study. For reference, the mean UPDRS Part III scores of PD subjects in two recent studies which focused on early PD detection were 16.9 [24] and 20.5 [4], respectively. A score of 8.4 is typical of patients with very mild disease severity. The control group included 67 healthy subjects without any sign of PD. The two groups were matched in terms of gender, age, education level, and years of experience with smartphones. All subjects were native Chinese speakers, 45 years or older. Table 2 summarizes the demographic and clinical information of the participants. We only included subjects who reported that they had used a touchscreen-equipped smartphone for more than one year, which can avoid or greatly reduce the confounding effect introduced by the variance in familiarity with common touch gestures. PD patients with cognitive impairment were excluded for the study.

Table 2: Summary of the demographic and clinical information of the 102 participants in this study.

	Early PDs	Controls	Sig
n (total n = 102)	35	67	n/a
Demographics			
Women #(%)	22(63%)	40(59.7%)	n/a
Men #(%)	13(37%)	27(40.3%)	n/a
Age yrs(std)	64.8(11.1)	62.7(10.0)	p=0.30
Education yrs(std)	9.3(3.3)	9.0(3.9)	p=0.44
Smartphone yrs(std)	3.4(1.6)	2.8(2.6)	p=0.15
Clinical Characteristics			
Disease Onset yrs(std)	2.3(1.9)	0(0)	p<0.001
UPDRS Part III avg(std)	8.4(3.7)	0(0)	p<0.001
Hoehn-Yahr avg(std)	1.3(0.8)	0(0)	p<0.001

The two groups are reasonable matched in terms of demographics as no significant differences are observed (two-sided Mann-Whitney U test).

The PD subjects were recruited from the PD clinic at Peking Union Medical College Hospital (Beijing). The study protocol (including both clinical evaluations and performing touch gestural interactions with a smartphone) was conducted during a single morning visit of each subject. All PD subjects were tested during the "ON" state[3] after their morning dose of the symptomatic relief medication.

Both the recruitment and the experimental procedures were approved by the Institutional Review Board of the hospital where the study was conducted. Subjects gave informed consent before their participation. They were told that they can quit the study at any time without providing any justification. The study was carried out according to institutional guidelines on research involving adult human beings. Figure 3 shows some sample pictures taken during the study.

[2]The Hoehn-Yahr scale is a commonly used system for describing how PD symptoms progress, ranging from 1 to 5. While control subjects did not exhibit any PD sign or motor abnormalities, they were rated as stage 0.

[3]The effectiveness of PD medication (e.g., levodopa) may fluctuate over time, a person can cycle through phases with good response to medication and reduced PD symptoms ("ON" state), and phases with poor response to medication and significant PD symptoms ("OFF" state).

Figure 3: Sample pictures of subjects in this study.

Procedure

Participants used our custom Android app on a Huawei P9 Plus smartphone running Android 7.0 for data collection. It has a 5.5 inch, 1080*1920 pixels, and pressure sensor-equipped touchscreen. The app has five major interfaces: a photo gallery interface to collect flick and pinch data, a pattern password interface to collect drag data, a Chinese handwriting input interface to collect handwriting data, an onscreen Qwerty-keyboard typing interface to collect typing data, and a classic AFT interface [6]. Navigation between the interfaces happened automatically.

At the beginning, a research staff member explained the whole procedure as well as each interface in the data collection app, and answered participants' questions. Participants were allowed to explore the app freely until they stated explicitly that they were ready to start the data collection process. Data acquired during this session was not included in the analysis.

Following this, the participant started with browsing pictures with a photo gallery interface by performing flick gestures. She/he performed left flicks for ten times to browse all pictures, and performed right flicks for another ten times to get back to the first picture. We did not put any constraint on participants and asked them to mimic their daily usage patterns as much as they could. Then the app automatically navigated to the pinch interface where the participant was asked to perform pinch gestures to zoom in the pictures for ten times, and to perform spread gestures to zoom out the pictures for another ten times. Then the app navigated to the pattern password interface which mimics the 3x3 lock screen of many Android devices. Each cell had a moderate diameter, and a digit was only triggered if the user's finger passed over the region of that cell. For each trial, the password was randomly generated with varying complexity (e.g., the number of finger direction change ranging from 1 to 4). We included such variety to guarantee that the detection was robust against pattern shapes and levels of complexity. The participant could see the correct pattern at the top of the screen and performed this task for ten times. Then the app repeated the above process for another two rounds. In total, we collected 60 flick samples (3 rounds x 2 types x 10 trials), 60 pinch samples (3 rounds x 2 types x 10 trials), and 30 drag samples (3 rounds x 10 trials) from each participant.

Then the participant used a custom Chinese handwriting input interface to input short Chinese excerpts for five minutes. After that, the participant typed with the Qwerty touch keyboard for another five minutes. The excerpts were selected from a Chinese short-text conversation corpus [55] based on the decreasing frequency order. Lastly, the participant performed the AFT test and completed the UPDRS Part III evaluations with a specialized neurologist. On average, it took 1.5 hours for each participant to complete the study.

To reduce the confounding effect of fatigue, the data collection app paused automatically when navigating between different interfaces to allow subjects to have a rest. In addition, they can pause at any time during the study as often as they like. Participants were instructed to sit comfortably and to operate the smartphone with their own styles. The experimenter explained the interfaces but did not mention the specific metrics that were being recorded. The goal was to simulate real-life and natural interactions with smartphones. Our app recorded the precise touch positions, the sequence of pressure values, the raw IMU outputs, and their corresponding times stamps of each touch gesture performed.

4 RESULTS

First, we investigate the discriminative power of the proposed features. Next, we assess the classification performance of using different categories of features (e.g., individual category or combined). We also compare the classification results between the proposed approach with external reference methods (i.e. the AFT test and the typing-based method).

Feature Exploration

Figure 4 presents the results of group-level (early PD patients vs. controls) statistical comparisons (two-sided Mann-Whitney U test) with respect to each individual feature, where we highlight the p-values of the null hypothesis tests that PD and control subjects come from the same population. It is evident that certain features differ significantly between groups. Specifically, 40 out of 46 features exhibit significant differences on at least one type of touch gestures.

To further investigate their discriminative power, we employed a feature selection approach suggested by Chen et al. [10], where we calculated the f-score for each feature.

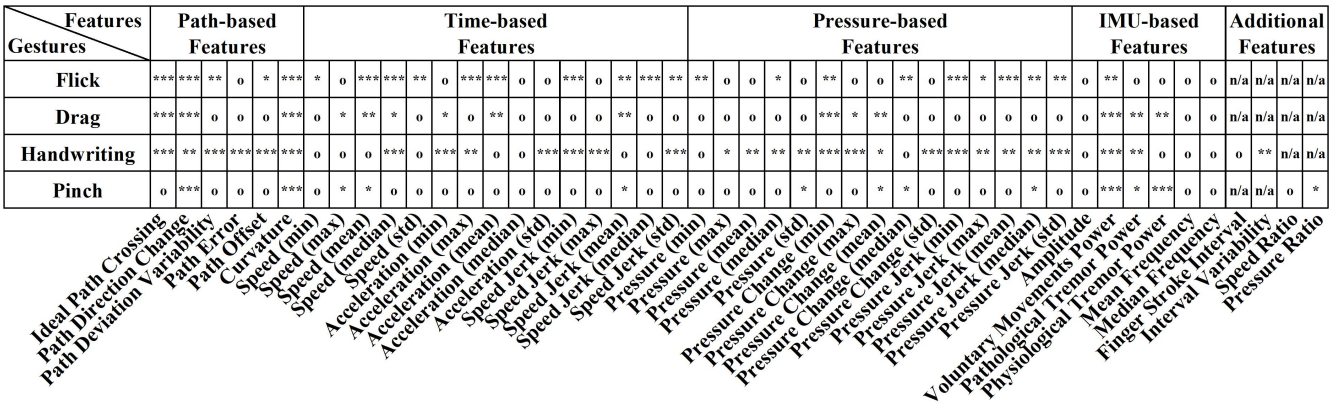

Figure 4: Results of statistical comparisons of each feature between early PD and control subjects. Significance is computed using the non-parametric two-sided Mann-Whitney U test and noted as: p<0.001(***), p<0.01(**), p<0.05(*), and p>0.05(o).

F-score can measure the discrimination of two sets of data samples. The f-score of the i^{th} feature is calculated as:

$$f(i) = \frac{(\bar{x}_i^{(+)} - \bar{x}_i)^2 + (\bar{x}_i^{(-)} - \bar{x}_i)^2}{\frac{1}{n_+ - 1}\sum (x_{k,i}^{(+)} - \bar{x}_i^{(+)})^2 + \frac{1}{n_- - 1}\sum (x_{k,i}^{(-)} - \bar{x}_i^{(-)})^2} \quad (1)$$

where $\bar{x}_i$ is the average of the i^{th} feature among all samples, $\bar{x}_i^{(+)}$ and $\bar{x}_i^{(-)}$ are the average values of the positive and negative samples, respectively. $x_{k,i}^{(+)}$ is the i^{th} feature value of the k^{th} positive sample while $x_{k,i}^{(-)}$ is the i^{th} feature value of the k^{th} negative sample. n_+ and n_- denote the number of positive and negative samples. The numerator indicates the discrimination between positive and negative sets, whereas the denominator indicates the discrimination within these sets. A higher f-score means that the feature is more likely to be discriminative when being used for classification. We calculated the f-score of each feature on each gesture type individually and ranked them based on the averaged value in decreasing order. Figure 5 shows the distributions of the most discriminant features in each category via box plots, which showed the upper and lower quartiles, as well as the median values. In general, as compared to healthy controls, PD patients exhibited less-efficient finger trajectories, milder and less stable speed and pressure, higher pathological tremor power, and lower voluntary movement power.

Two path-based features (curvature and path direction change) exhibited significant differences on all types of touch gestures, thereby their discriminative power was highlighted. PD patients exhibited on average larger numbers of movement direction change (normalized based on stroke length, flick: 0.10 vs. 0.02, p<0.001; drag: 0.18 vs. 0.10, p<0.001; handwriting: 0.35 vs. 0.21, p<0.001; pinch: 0.27 vs. 0.16, p<0.001) and higher curvature values (flick: 0.18 vs. 0.08, p<0.001; drag: 0.37 vs. 0.29, p<0.001; handwriting: 0.74 vs. 0.68, p=0.003; pinch: 0.41 vs. 0.26, p<0.001). These features can reflect the directness of finger trajectories, which may further indicate the efficiency when making fine finger movements. The results confirmed our hypothesis that motor impairments in PD can affect the efficiency in term of finger trajectories.

As derived from the distributions of the most discriminative time-based features, PD patients exhibited on average lower values of maximum finger velocities when performing all the four gestures, although statistical significances were achieved only on drag (0.18 vs. 0.26, p=0.018) and pinch gestures (0.09 vs. 0.18, p=0.027). We attribute such results to the effects of motor symptoms, especially bradykinesia (slowness of movement), which may low down finger reflexes causing PD patients to make relatively slower finger movements. These results were also consistent with prior findings the PD patients had longer key holding time (HT) during typing activities [3, 24]. At the same time, we also found that PD patients had higher absolute acceleration values when performing flick (0.60 vs. 0.46, p<0.001) and drag (0.47 vs. 0.42, p<0.002) gestures, which indicated unstable swiping speed when interacting with touchscreens. This may be caused by the rigidity (muscle stiffness) of PD patients which can disturb the intended operating behaviors.

In our experiment, early PD subjects and controls also exhibited significant differences in terms of pressure applied to perform certain touchscreen gestures. Specifically, PD subjects produced significantly lower values in median pressure than control subjects when performing flick (0.90 vs. 0.95, p=0.014) and handwriting (0.88 vs. 0.92, p=0.004) gestures. At the same time, PD patients exhibited significantly higher values in pressure change when performing flick (0.15 vs. 0.09, p=0.006), drag (0.18 vs. 0.11, p<0.001), and handwriting (0.16 vs. 0.07, p<0.001) gestures. The former could be attributed to inadequately-scaled movements in term of amplitude that constitute manifestations of hypokinesia that is present

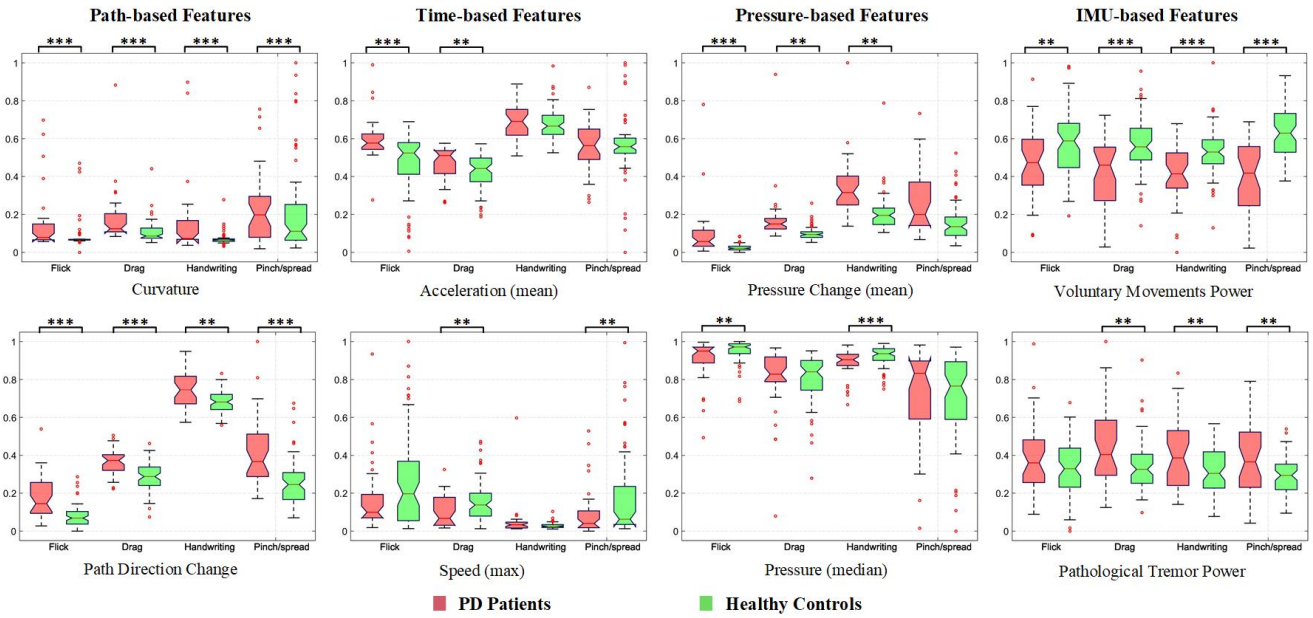

Figure 5: Group-wise comparisons on the most discriminative features of each feature category. Boxplots represent the feature distributions of 35 PD and 67 control subjects. Statistical significance is noted as: p<0.001(***), p<0.01(**), and p<0.05(*).

in PD, which has been reported by related prior research [24]. The latter confirmed a recent clinical finding that the variance of contact pressure increases as the severity of the associated motor symptom increases [50].

Besides, PD subjects exhibited significantly lower voluntary movement power (spectrum power on frequency band of 1-3.5 Hz [20] when performing all four gestures, which can result from the slowness of intentional movements. At the same time, PD subjects produced significantly higher pathological tremor power (spectrum power on frequency band of 3.5-7.5 Hz) when performing drag (0.45 vs. 0.34, p=0.005), handwriting (0.41 vs. 0.31, p=0.008), and pinch (0.39 vs. 0.29, p=0.017) gestures. Although consistent results were achieved on flick gestures (0.39 vs. 0.33, p=0.278), the difference was not statistically significant. These results were consistent with findings in prior studies investigating the action or postural tremor of PD [7, 18, 20, 47], which confirmed that the IMU data collected during touchscreen gestural interactions can also be informative and useful for PD assessment.

Classification Results

We explored a set of common classification methods including logistic regression, support vector machine (SVM), random forest, and AdaBoost in our preliminary studies, and achieved best results with linear kernel SVMs (implemented by the LIBSVM toolkit [9]). Therefore, we only report the SVM results in the rest of paper due to space limit. We employed multiple tests and metrics to measure the classification performance, including the Receiver Operating Characteristic (ROC) analysis, sensitivity (i.e. true positive rate, the percentage of PD patients who are correctly identified as having PD), specificity (i.e. true negative rate, the percentage of healthy subjects who are correctly identified as being healthy), and the F1 score (i.e. the harmonic mean of precision and sensitivity). All these metrics allow a reliable comparison of classification performance, even when the class size is not fully balanced, as it is the case of our dataset (35 PD, 67 controls).

The ROC curve is created by plotting true positive rate (sensitivity) against false positive rate (1-specificity) at various threshold settings[13]. The area under the curve (AUC) can be interpreted as the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one (assuming 'positive' ranks higher than 'negative'). A test with no better accuracy than chance has an AUC of 0.5 while a test with perfect accuracy has an AUC of 1. A sampling with replacement method (1000 bootstraps) defines a ROC distribution from which we compute the average and confidence intervals of the AUC values. Moreover, we estimated the sensitivity and specificity values in our study via the closest-to-(0,1) criterion when defining the cut-off point [46]. We leveraged leave-one-subject-out-cross-validation. Specifically, for each pass, the data of 1

Table 3: Classification results of using different feature combinations.

Feature Set	AUC [5% 95%]	Sens	Spec	F1
A	0.90 [0.83 0.95]	0.83	0.82	0.76
B	0.88 [0.79 0.94]	0.80	0.84	0.76
C	0.86 [0.77 0.92]	0.80	0.78	0.72
D	0.89 [0.82 0.94]	0.80	0.85	0.77
A+B	0.91 [0.83 0.95]	0.83	0.85	0.78
A+C	0.92 [0.85 0.96]	0.80	0.87	0.78
A+D	0.91 [0.84 0.96]	0.83	0.85	0.78
B+C	0.92 [0.84 0.96]	0.83	0.81	0.75
B+D	0.93 [0.87 0.97]	0.86	0.85	0.80
C+D	0.93 [0.86 0.96]	0.77	0.87	0.76
A+B+C	0.92 [0.84 0.96]	0.86	0.82	0.78
A+B+D	0.93 [0.87 0.97]	0.80	0.91	0.81
A+C+D	0.94 [0.89 0.97]	0.83	0.88	0.81
B+C+D	0.93 [0.87 0.97]	0.89	0.88	0.84
A+B+C+D	0.95 [0.90 0.98]	0.89	0.88	0.84

A : Path-based Features, B : Time-based Features,
C : Pressure-based Features, D : IMU-based Features.

participant was used as the test set and the data of the remaining 101 subjects was used as the training set. Therefore, the results reported follow were user-independent.

We first investigated the classification performance of using each individual category of features, as well as using all possible combinations of feature categories. Such analysis can provide insights on how well PD and control subjects can be discriminated by each category of features, and how different categories of features can complement one another. Table 3 summarized the results achieved on the whole dataset including all four types of touch gestures. We report average AUC for the bootstrapped ROC distribution, AUC confidence interval computed as the [5%, 95%] percentiles on the resulting AUC values, sensitivity, specificity, and F1 scores. We achieved promising results even with only one category of features. Specifically, we achieved an AUC of 0.90 [0.83 0.95] with path-based features, 0.88 [0.79 0.94] with time-based features, 0.86 [0.77 0.92] with pressure-based features, and 0.89 [0.82 0.94] with IMU-based features, which highlighted their discriminatory power. Not surprisingly, the classification performances increased when using multiple categories of features, though some combinations led to more improvements than others. Using all of the features together led to the highest AUC of 0.95 [0.90 0.98] with the 0.89/0.88 sensitivity/specificity.

We also compared the classification results of employing each touch gesture solely, using the combination of all gestures, with external baseline methods including the AFT test and the typing-based method proposed by Iakovakis et al. [24]. We replicate the evaluation framework used in our methods to test the classification performance of these two reference metrics in our cohort. Table 4 summarizes the results. The AFT test achieved an AUC of 0.83 [0.75 0.90] and the typing-based method achieved an AUC of 0.88 [0.80 0.93]. These results were very close to what prior studies reported (e.g., AFT achieved an AUC of 0.85 in [3], typing achieved an AUC of 0.88 in [3] and 0.92 in [24]). For touch gestures, we found that each gesture performed well on its own. Each gesture solely can achieve comparable performance with the typing-based method, and can outperform the AFT test. When combined, touch gestures achieved an AUC of 0.95 [0.90 0.98], which outperformed the typing-based method which presented an AUC of 0.88 [0.80 0.93] in our dataset. At the same time, we also found that the combination of

Table 4: Classification results of using different gestures, compared with external reference methods.

Tasks	AUC [5% 95%]	Sens	Spec	F1
Flick(F)	0.88 [0.80 0.94]	0.77	0.75	0.68
Drag(D)	0.92 [0.84 0.96]	0.86	0.88	0.82
Handwriting(H)	0.89 [0.82 0.94]	0.71	0.85	0.71
Pinch(P)	0.92 [0.85 0.96]	0.80	0.82	0.75
Gestures(F+D+H+P)	0.95 [0.90 0.98]	0.89	0.88	0.84
Typing	0.88 [0.80 0.93]	0.74	0.85	0.73
Typing+Gestures	0.97 [0.92 0.99]	0.89	0.90	0.85
AFT	0.83 [0.75 0.90]	0.86	0.69	0.70

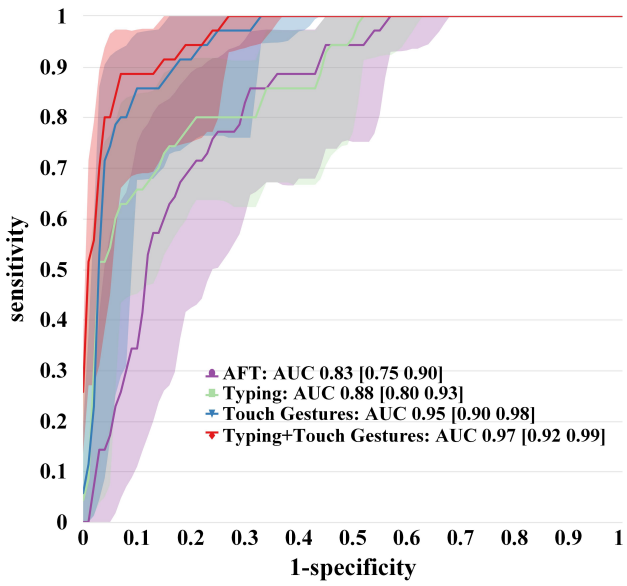

Figure 6: ROC curves demonstrating the classification performance of the proposed approach and exiting methods.

touch gestures and typing led to better performance (AUC of 0.97 [0.92 0.99]) than leveraging either class solely, which indicated that they can complement well one the other well. Figure 6 shows the comparison of ROC curves achieved by different methods. The improvement achieved in the classification result, compared to the AFT test which is a commonly used clinical method to quantify upper limbs dexterity, may due to the fact that our features have been carefully defined to capture the motor abnormalities that are a direct representation of PD signs, including tremor, slowness, and rigidity.

5 DISCUSSION

This work explored detecting motor impairment in early PD from common touch gestural interactions. We contribute a set of useful empirical findings for both researchers and clinical practitioners. In general, compared to healthy controls, our results indicated that early PD subjects exhibited less-efficient finger trajectories (e.g., significantly higher values in path curvature and direction change), milder and less stable speed and pressure, higher tremor power, and lower voluntary movement power when they performed touch gestures on smartphones. This confirmed our hypothesis that by analyzing data from rich sensors such as IMUs, pressure sensors and touchscreens, we can detect the abnormalities in touch interaction kinetics caused by PD signs such as tremor, bradykinesia, and rigidity. As expected, the best classification performance (0.95 AUC with 0.89/0.88 sensitivity/specificity) was achieved by using all categories of features together, which indicated they can complement each other to improve the detection performance.

Our results also demonstrated that each of four common touch gestures on its own can outperformed the clinical AFT test. We attribute the improvement to the larger variety of features that can better reflect PD signs. In comparison, the AFT test relied on time-based metrics solely. We also found touch gestures and typing complemented each other and improved classification performance when combined. When comparing between single touch gestures, we found that drag gestures (0.92 AUC with 0.86/0.88 sensitivity/specificity) outperformed others on almost all metrics. This may due to the fact that users were more constrained when performing drag gestures in the current experimental setting, where they had specific paths to trace, which included explicit starting point, end point, and the crossing goals in between. Such constraints may limit the interfering effect of personal style when users perform such gestures, especially compared with performing more flexible gestures such as flicks.

Our study was conducted in a controlled environment. Although participants were asked to perform these gestures as they would normally do to mimic actual routine use of smartphones, further longitudinal and large-scale deployment is necessary to discard the impact of controlled study on participants' interaction behaviors, where data will be captured unobtrusively and implicitly from users' naturalistic and daily smartphone interactions. Furthermore, it is worth exploration of developing a regression model based on the current classification model, which could make it possible to conduct fine-grained evaluation on the motor function and the disease progression. Addition information from user's everyday smartphone interaction, such as voice signal during phone calls, IMU data during idle periods, could be used to complement our analysis.

Privacy issues should be considered since the proposed approach relies on the data collected unobtrusively and implicitly from routine smartphone interactions. Actually, this approach is privacy-aware, as it is based on the low-level features calculated from raw sensor outputs, without requiring the actual content and high-level information. However, the results of such approach indicate users' health status which belongs to the category of sensitive personal data. Future applications should ensure that the users are aware of what data would be collected and when it would be collected. They should comply with ethical guidelines and data protection regulations as well.

6 CONCLUSION

Our investigation over 102 subjects' data showed that it was practical and promising to detect motor impairment in early PD via sensing and analyzing users' common gestural interactions on smartphones. Our machine learning based approach achieved an AUC of 0.95 and 0.89/0.88 sensitivity/specificity in discriminating early PD subjects from age-matched healthy controls. This approach also outperformed the alternating finger tapping (AFT) test, a well-established PD motor test in clinical settings. Our investigation also revealed the discriminative power of both the features and the gestures.

A large amount of data were acquired for this work. In order for other researchers to replicate and advance our results, we release our 102-subject (35 early PD, 67 healthy controls) dataset under a BSD license (please contact the corresponding author to make a request). It can serve as a common ground for future research and performance comparison of different approaches.

ACKNOWLEDGMENTS

We thank participants in our study, and thank reviewers for their constructive feedback. This work was supported by National Key R&D Program of China (Grant No. 2016YFB1001405), National Natural Science Foundation of China (Grant No. 61872349), Key Research Program of Frontier Sciences, CAS (Grant No. QYZDY-SSW-JSC041), and the CAS Pioneer Hundred Talents Program.

REFERENCES

[1] Stephen A Antos, Mark V Albert, and Konrad P Kording. 2014. Hand, belt, pocket or bag: Practical activity tracking with mobile phones. *Journal of Neuroscience Methods* 231 (2014), 22–30.

[2] Siddharth Arora, Vinayak Venkataraman, Andong Zhan, Sean R Donohue, Kevin M Biglan, E R Dorsey, and Max A Little. 2015. Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Parkinsonism and Related Disorders* 21, 6 (2015), 650–653.

[3] Teresa Arroyo-Gallego, M Ledesma-Carbayo, Alvaro Sánchez-Ferro, Ian Butterworth, Carlos S Mendoza, Michele Matarazzo, Paloma Montero, Roberto López-Blanco, Verónica Puertas-Martín, Rocıo Trincado, et al. 2017. Detection of motor impairment in Parkinson's disease via mobile touchscreen typing. *IEEE Trans Biomed Eng* 64, 9 (2017), 1994–2002.

[4] Teresa Arroyo-Gallego, María J Ledesma-Carbayo, Ian Butterworth, Michele Matarazzo, Paloma Montero-Escribano, Verónica Puertas-Martín, Martha L Gray, Luca Giancardo, and Álvaro Sánchez-Ferro. 2018. Detecting Motor Impairment in Early Parkinson's Disease via Natural Typing Interaction With Keyboards: Validation of the neuroQWERTY Approach in an Uncontrolled At-Home Setting. *Journal of medical Internet research* 20, 3 (2018).

[5] Jeremy N Bailenson, Nick Yee, Scott Brave, Dan Merget, and David Koslow. 2007. Virtual interpersonal touch: expressing and recognizing emotions through haptic devices. *Human-Computer Interaction* 22, 3 (2007), 325–353.

[6] Brian M Bot, Christine Suver, Elias Chaibub Neto, Michael R Kellen, Arno Klein, Christopher Bare, Megan Doerr, Abhishek Pratap, John Wilbanks, E Ray Dorsey, et al. 2016. The mPower study, Parkinson disease mobile data collected using ResearchKit. *Scientific Data* 3 (2016), 160011–160011.

[7] Filipa L Campos, Miguel Carvalho, Ana Clara Cristovao, Goun Je, Graca Baltazar, Antonio J Salgado, Yoonseong Kim, and Nuno Sousa. 2013. Rodent models of Parkinson's disease: beyond the motor symptomatology. *Frontiers in Behavioral Neuroscience* 7, 175 (2013), 175–175.

[8] Benoit Carignan, Jean-François Daneault, and Christian Duval. 2015. Measuring tremor with a smartphone. In *Mobile Health Technologies*. Springer, 359–374.

[9] Chihchung Chang and Chihjen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011), 27.

[10] Yi-Wei Chen and Chih-Jen Lin. 2006. Combining SVMs with various feature selection strategies. In *Feature extraction*. Springer, 315–324.

[11] Lonneke M L De Lau and Monique M B Breteler. 2006. Epidemiology of Parkinson's disease. *Lancet Neurology* 5, 6 (2006), 525–535.

[12] Robert J Ellis, Yee Sien Ng, Shenggao Zhu, Dawn M Tan, Boyd Anderson, Gottfried Schlaug, and Ye Wang. 2015. A validated smartphone-based assessment of gait and gait variability in Parkinson's disease. *PLoS one* 10, 10 (2015), e0141694.

[13] Tom Fawcett. 2005. An introduction to ROC analysis. *Pattern Recognition Letters* 27, 8 (2005), 861–874.

[14] Moshe Gabel, Ran Gilad-Bachrach, Erin Renshaw, and Assaf Schuster. 2012. Full body gait analysis with Kinect. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 1964–1967.

[15] Yuan Gao, Nadia Bianchi-Berthouze, and Hongying Meng. 2012. What does touch tell us about emotions in touchscreen-based gameplay? *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 4 (2012), 31.

[16] Joseph Ghika, Allen W Wiegner, Jian Jun Fang, Llewelyn Davies, Robert R Young, and John H Growdon. 1993. Portable system for quantifying motor abnormalities in Parkinson's disease. *IEEE Transactions on Biomedical Engineering* 40, 3 (1993), 276–283.

[17] Christopher G Goetz, Stanley Fahn, Pablo Martinez-Martin, Werner Poewe, Cristina Sampaio, Glenn T Stebbins, Matthew B Stern, Barbara C Tilley, Richard Dodel, Bruno Dubois, et al. 2007. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): process, format, and clinimetric testing plan. *Movement Disorders* 22, 1 (2007), 41–47.

[18] Robert I Griffiths, Katya Kotschet, Sian Arfon, Zheng Ming Xu, William Johnson, John Drago, Andrew Evans, Peter Kempster, Sanjay Raghav, and Malcolm K Horne. 2012. Automated assessment of bradykinesia and dyskinesia in Parkinson's disease. *Journal of Parkinson's disease* 2, 1 (2012), 47–55.

[19] Peter Hagell, Sara Nordling, Jan Reimer, Martin Grabowski, and Ulf Persson. 2002. Resource use and costs in a Swedish cohort of patients with Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society* 17, 6 (2002), 1213–1220.

[20] Tjitske Heida, Eva Christine Wentink, and Enrico Marani. 2013. Power spectral density analysis of physiological, rest and action tremor in Parkinson's disease patients treated with deep brain stimulation. *Journal of neuroengineering and rehabilitation* 10, 1 (2013), 70.

[21] Matthew J Hertenstein, Rachel Holmes, Margaret McCullough, and Dacher Keltner. 2009. The communication of emotion via touch. *Emotion* 9, 4 (2009), 566.

[22] D Hirtz, DJ Thurman, K Gwinn-Hardy, M Mohamed, AR Chaudhuri, and R Zalutsky. 2007. How common are the "common" neurologic disorders? *Neurology* 68, 5 (2007), 326–337.

[23] J D Carl Nikolaus Homann, Klaudia Suppan, Karoline Wenzel, Gavin Giovannoni, Gert Ivanic, Susanne Horner, Erwin Ott, and Hanspeter Hartung. 2000. The Bradykinesia Akinesia Incoordination Test (BRAIN TEST), an objective and user-friendly means to evaluate patients with parkinsonism. *Movement Disorders* 15, 4 (2000), 641–647.

[24] 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), 7663.

[25] Preeti Khanna and M Sasikumar. 2010. Recognising Emotions from Keyboard Stroke Pattern. *International Journal of Computer Applications* 11, 9 (2010), 24–28.

[26] Jurgen Konczak, Hermann Ackermann, Ingo Hertrich, Sybille Spieker, and Johannes Dichgans. 1997. Control of repetitive lip and finger movements in parkinson's disease: Influence of external timing signals and simultaneous execution on motor performance. *Mov Disord* 12, 5 (1997), 665–676.

[27] N Kostikis, Dimitris Hristuvarsakelis, M Arnaoutoglou, and C Kotsavasiloglou. 2015. A Smartphone-Based Tool for Assessing Parkinsonian Hand Tremor. *IEEE Journal of Biomedical and Health Informatics* 19, 6 (2015), 1835–1842.

[28] Kun-Chan Lan and Wen-Yuah Shih. 2014. Early Diagnosis of Parkinson's Disease Using a Smartphone. *Procedia Computer Science* 34 (2014), 305–312.

[29] Chae Young Lee, Seong Jun Kang, Sangkyoon Hong, Hyeoil Ma, Unjoo Lee, and Yun Joong Kim. 2016. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson's Disease. *PLOS ONE* 11, 7 (2016).

[30] Robert LeMoyne and Timothy Mastroianni. 2015. Use of smartphones and portable media devices for quantifying human movement characteristics of gait, tendon reflex response, and Parkinson's disease hand tremor. In *Mobile Health Technologies*. Springer, 335–358.

[31] Hairong Lv, Zhonglin Lin, Wen Jun Yin, and Jin Dong. 2008. Emotion recognition based on pressure sensor keyboards. (2008), 1089–1092.

[32] I Scott Mackenzie, Tatu Kauppinen, and Miika Silfverberg. 2001. Accuracy measures for evaluating computer pointing devices. (2001), 9–16.

[33] Alexander Mariakakis, Sayna Parsi, Shwetak N Patel, and Jacob O Wobbrock. 2018. Drunk User Interfaces: Determining Blood Alcohol Level through Everyday Smartphone Tasks. (2018), 234.

[34] Daniel Martín-Albo, Luis A Leiva, and Réjean Plamondon. 2016. On the design of personal digital bodyguards: Impact of hardware resolution on handwriting analysis. In *Frontiers in Handwriting Recognition (ICFHR), 2016 15th International Conference on*. IEEE, 174–179.

[35] Pietro Mazzoni, Britne Shabbott, and Juan Camilo Cortes. 2012. Motor Control Abnormalities in Parkinson's Disease. *Cold Spring Harbor Perspectives in Medicine* 2, 6 (2012).

[36] Roisin Mcnaney, Ivan Poliakov, John Vines, Madeline Balaam, Pengfei Zhang, and Patrick Olivier. 2015. LApp: A Speech Loudness Application for People with Parkinson's on Google Glass. (2015), 497–500.

[37] Roisin Mcnaney, John Vines, Daniel Roggen, Madeline Balaam, Pengfei Zhang, Ivan Poliakov, and Patrick Olivier. 2014. Exploring the acceptability of google glass as an everyday assistive device for people with parkinson's. (2014), 2551–2554.

[38] Philipp Mock, Peter Gerjets, Maike Tibus, Ulrich Trautwein, Korbinian Möller, and Wolfgang Rosenstiel. 2016. Using touchscreen interaction data to predict cognitive workload. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ACM, 349–356.

[39] Torin Monahan and Jill A Fisher. 2010. Benefits of 'observer effects': lessons from the field. *Qualitative Research* 10, 3 (2010), 357–376.

[40] Daniel L Murman. 2012. Early treatment of Parkinson's disease: opportunities for managed care. *The American Journal of Managed Care* 18 (2012).

[41] Alastair J Noyce, A Nagy, Shami Acharya, Shahrzad Hadavi, Jonathan P Bestwick, Julian Fearnley, Andrew J Lees, and Gavin Giovannoni. 2014. Bradykinesia-akinesia incoordination test: validating an online keyboard test of upper limb function. *PLOS ONE* 9, 4 (2014).

[42] Fernando L Pagan. 2012. Improving Outcomes Through Early Diagnosis of Parkinson's Disease. *The American Journal of Managed Care* 18 (2012).

[43] Luca Palmerini, Sabato Mellone, Laura Rocchi, and Lorenzo Chiari. 2011. Dimensionality reduction for the quantitative evaluation of a smartphone-based Timed Up and Go test. 2011 (2011), 7179–7182.

[44] Di Pan, Rohit Dhall, Abraham Lieberman, and Diana B Petitti. 2015. A Mobile Cloud-Based Parkinson's Disease Assessment System for Home-Based Monitoring. *Jmir mhealth and uhealth* 3, 1 (2015).

[45] V Parra, G Figueras, M Huerta, A Marzinotto, R Gonzalez, and R Alvizu. 2013. A smartphone application for parkinson tremor detection. In *Proc. of the IEEE Latin-American Conference on Communications*.

[46] Neil J Perkins and Enrique F Schisterman. 2006. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *American journal of epidemiology* 163, 7 (2006), 670–675.

[47] Shyam V. Perumal and Ravi Sankar. 2016. Gait and tremor assessment for patients with Parkinson's disease using wearable sensors. *Ict Express* 2, 4 (2016), 21–24.

[48] Rejean Plamondon, Giuseppe Pirlo, Eric Anquetil, Céline Rémi, Hans-Leo Teulings, and Masaki Nakagawa. 2018. Personal digital bodyguards for e-security, e-learning and e-health: A prospective survey. *Pattern Recognition* 81 (2018), 633–659.

[49] Blake P Printy, Lindsey M Renken, John P Herrmann, Isac Lee, Bryant Johnson, Emily Knight, Georgeta Varga, and Diane Whitmer. 2014. Smartphone application for classification of motor impairment severity in Parkinson's disease. 2014 (2014), 2686–2689.

[50] Matthew P Rearick, George E Stelmach, Berta Leis, and Marco Santello. 2002. Coordination and control of forces during multifingered grasping in Parkinson's disease. *Experimental neurology* 177, 2 (2002), 428–442.

[51] K Sajith, Vikas Darade, and Subhasis Chaudhuri. 2013. Hand Tremor Analysis Using Rigid Body Manipulation in a Dynamic Virtual Haptic Environment. In *Proceedings of Conference on Advances In Robotics*. ACM, 1–5.

[52] Guarionex Salivia and Juan Pablo Hourcade. 2011. Identification of pointing difficulties of two individuals with Parkinson's disease via a sub-movement analysis. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM, 137–140.

[53] Alvaro Sanchezferro, Morad Elshehabi, Catarina Godinho, Dina Salkovic, Markus A Hobert, Josefa Domingos, Janet M T Van Uem, Joaquim J Ferreira, and Walter Maetzler. 2016. New methods for the assessment of Parkinson's disease (2005 to 2015): A systematic review. *Movement Disorders* 31, 9 (2016), 1283–1292.

[54] Anette Schrag, Laura Horsfall, Kate Walters, Alastair J Noyce, and Irene Petersen. 2015. Prediagnostic presentations of Parkinson's disease in primary care: a case-control study. *Lancet Neurology* 14, 1 (2015), 57–64.

[55] Lifeng Shang, Zhengdong Lu, and Hang Li. 2015. Neural Responding Machine for Short-Text Conversation. (2015), 52–58.

[56] Caroline M Tanner and D A Aston. 2000. Epidemiology of Parkinson's disease and akinetic syndromes. *Current Opinion in Neurology* 13, 4 (2000), 427–430.

[57] Ana Lisa Taylor Tavares, Gregory S X E Jefferis, Mandy Miller Koop, Bruce C Hill, Trevor Hastie, Gary Heit, and Helen Brontestewart. 2005. Quantitative measurements of alternating finger tapping in Parkinson's disease correlate with UPDRS motor disability and reveal the improvement in fine motor control from medication and deep brain stimulation. *Movement Disorders* 20, 10 (2005), 1286–1298.

[58] H L Teulings, Jose L Contrerasvidal, George E Stelmach, and Charles H Adler. 1997. Parkinsonism Reduces Coordination of Fingers, Wrist, and Arm in Fine Motor Control. *Experimental Neurology* 146, 1 (1997), 159–170.

[59] John A Waterston, Graham R Barnes, Madeleine Grealy, and Sue Collins. 1996. Abnormalities of smooth eye and head movement control in Parkinson's disease. *Annals of Neurology* 39, 6 (1996), 749–760.