Towards unobtrusive Parkinson's disease detection via motor symptoms severity inference from multimodal smartphone-sensor data

Dimitrios Iakovakis[1], Stelios Hadjidimitriou[1], Vasileios Charisis[1], Konstantinos Kyritsis[1], Alexandros Papadopoulos[1], Michael Stadtschnitzer[2], Hagen Jaeger[2], Ioannis Dagklis[3], Sevasti Bostantjopoulou[3], Zoe Katsarou[4], Lisa Klingelhoefer[5], Simone Mayer[5], Heinz Reichmann[5], Dhaval Trivedi[6], Aleksandra Podlewska[6], Alexandra Rizos[6], Karrol Ray Chaudhuri[6], Anastasios Delopoulos[1] and Leontios J. Hadjileontiadis[1,7]

1 Department of Electrical & Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece 2 Fraunhofer IAIS, Schloß Birlinghoven, Sankt Augustin, Germany 3 G. Papanikolaou Hospital, 3rd Neurological Clinic, Thessaloniki, Greece 4 Department of Neurology, Hippokration Hospital, Thessaloniki, Greece 5 Department of Neurology, Technical University Dresden, Dresden, Germany 6 King's College Hospital NHS Foundation Trust, London, United Kingdom 7 Khalifa University of Science and Technology, Abu Dhabi, UAE.

Objective

To provide clinically-corroborated evidence of the Parkinson's disease (PD) diagnostic potential of machine learning-based approaches for motor symptoms severity inference via multimodal data, passively captured during the natural use of smartphones.

Background

PD symptoms can be mild in the early stages and they usually go unnoticed, leaving the disease undiagnosed for years [1]. Subtle motor manifestations may start five to six years prior to PD clinical diagnosis and thereafter progress quickly [2]. Motor impairment affects daily activities and can severely impact patients' quality over the course of the disease. Information derived from mobile electronic sensors can provide, via algorithmic transformation, objective and dense information of an individual's motor status, allowing for frequent relevant symptoms early screening and subsequent monitoring.

Methodology

Table 1 Medically valid dataset description.

	Group		
	PD patients	Healthy controls	p-value
No. of subjects	47	23	0.075
Male/Female rate	66/33	43/57	57/43
Mean Age (std)	61.9 (7.6)	56.0 (10.9)	0.2
Mean UPRDS Part III (std)	18.0 (10.5)	2.70	p<0.01

We analyzed longitudinal recordings of tri-axial accelerometer, voice and keystroke timing data, captured passively from 70 PD patients and healthy controls (HC) [Table 1], in their daily life via the iPrognosis Android smartphone application, for relevant motor symptoms severity inference. [Fig1], [Fig2] and [Fig3] depict the proposed processing pipelines, whereas, [Table 2], [Table 3] and [Table 4] epitomize the datasets exploited per modality. Subjects underwent medical evaluation in order to obtain the UPDRS Part-III scores used as ground truth.

Touchscreen typing analysis pipeline

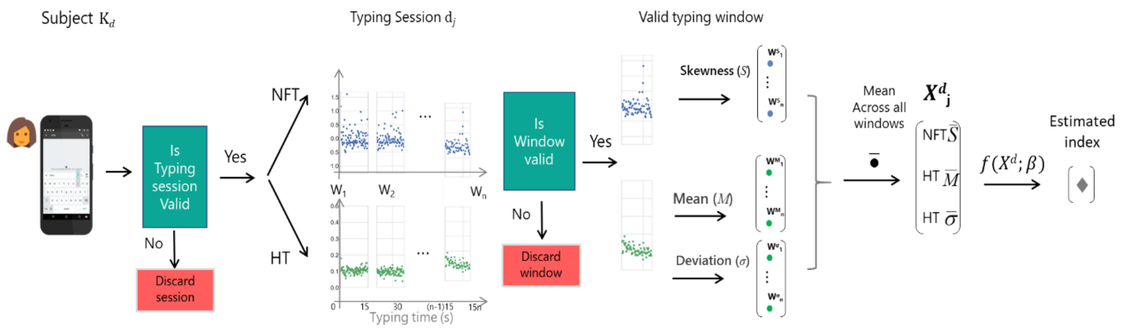

Fig 1. Processing pipeline of keystroke timing data of a subject's typing sessions to quantify the severity of rigidity and bradykinesia. Valid typing sessions are considered those with at least eight keystroke from where hold-time (HT) and intra-keystroke flight-time (FT) sequences are split in time-windows (with at least four keystrokes to considered valid), and thereafter statistical features are extracted from valid windows and used as input in an already trained machine learning model (f) which produces the estimated severity indices per typing session. The mean of the estimated indices over all typing sessions constitutes the aggregated index quantifying the severity of the rigidity and bradykinesia for the particular subject. Parameters were initiated on a separate dataset (D2) to estimate the correspondence UPDRS Part III single item scores under guided experimental scenarios in-the-clinic. See [3] for more details.

Table 2 Keystroke timing dataset. Recordings consist of the timing information of the typing activity while using the custom keyboard for the routine typing activities.

	Group		
	PD patients	Healthy controls	Overall
No. of subjects	10	8	18
Avg. duration of data contribution per subject, days (std)	25 (60)	19 (25)	22 (47)
Total no. of recordings	769	1,818	2,587

No.: number; Avg.: average; Std: standard deviation

Voice analysis pipeline

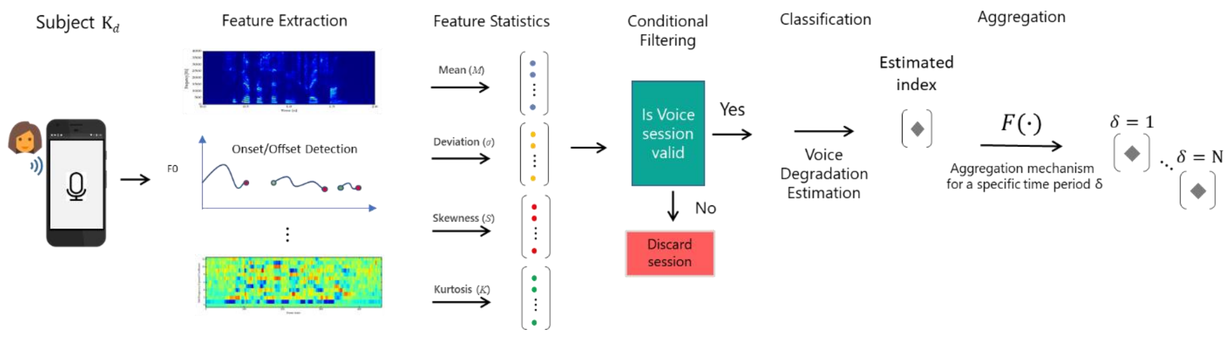

Fig 2. Processing pipeline of voice sessions to estimate the level of speech impairment. Voice features are based on the onset/offset detection of the fundamental frequency and subsequent spectral descriptor extraction for each recording, which after conditional filtering are fed into a machine learning model. The decisions are aggregated and averaged over a certain period of time, producing the estimator of speech impairment.

Table 3 Voice features dataset. Recordings consist of extracted features of the voice signals that were captured during phone calls.

	Group		
	PD patients	Healthy controls	Overall
No. of subjects	26	15	41
Avg. duration of data contribution per subject, days (std)	32 (44)	94 (72)	55 (63)
Total no. of recordings	376	664	1,040

No.: number; Avg.: average; Std: standard deviation

Tremor analysis pipeline

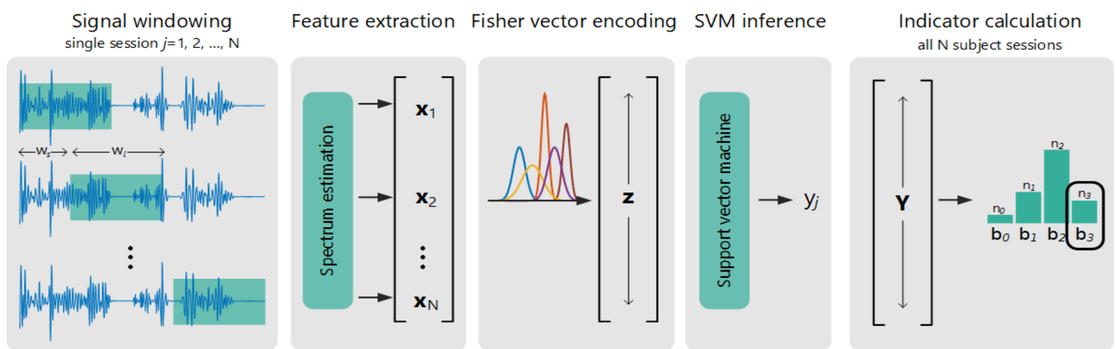

Fig 3. Five steps of the tremor index estimation pipeline using in-the-wild data. A single session is initially processed by segmenting it into frames using a sliding window. Next, spectrum estimation for each frame is performed using Welch's periodogram approach. All session features x are encoded into a Fisher Vector z of constant length. A pre-trained SVM is then used to provide the session score y. The SVM is trained on a lab-controlled dataset of 30 subjects involving both PD patients and healthy controls. Aggregation is performed by collecting the SVM scores from all the subject's sessions and form a normalized histogram with 4 bins [-inf, -0.5, +0.5, +inf]. The estimated index corresponds to the amount of sessions within the [+0.5, +inf] bin.

Table 4 Accelerometer dataset. Recordings consist of tri-axial accelerometer signals captured during phone calls.

	Group		
	PD patients	Healthy controls	Overall
No. of subjects	29	16	45
Avg. duration of data contribution per subject, days (std)	54 (82)	156 (95)	90 (100)
Total no. of recordings	1,717	2,875	4,592

No.: number; Avg.: average; Std: standard deviation

Results & Conclusions

Table 5 Diagnostic performance of the symptom-inference pipelines. Highest Area under the ROC curve were achieved via the rigidity (0.89) and bradykinesia (0.82) severity estimators followed by speech impairment (0.68) and tremor modalities (0.67).

Inferred symptom / Diagnostic Properties	Speech impairment	Action tremor	Rigidity	Bradykinesia
Area Under the ROC curve	0.68	0.67	0.89	0.82
Sensitivity	0.55	0.50	1.00	1.00
Specificity	0.72	0.88	0.80	0.70
Diagnostic Accuracy	0.67	0.77	0.89	0.83

The proposed methods for motor symptoms inference show promising PD diagnostic performance in our relatively small clinically-evaluated cohorts. Our results highlight the potential of evolving these methods into an objective PD screening/monitoring tool that could support clinical diagnosis, drug response assessment and decision-making. Passive capturing of the required input data further fosters evaluation of individuals' natural behavior, as well as long-term adherence.

References:1. Hawkes, C. H., Del Tredici, K., and Braak, H. (2010). A timeline for Parkinson's disease. Parkinsonism Relat. Disord. 16, 79–84. doi: 10.1016/j.parkreldis.2009.08.007
2. Fereshtehnejad, S. M., et al. "Evolution of Motor and Non-Motor Manifestations in Prodromal Parkinsonism." MOVEMENT DISORDERS. Vol. 33. No. 12. 111 RIVER ST, HOBOKEN 07030-5774, NJ USA: WILEY, 2018. Hawkes, Christopher H., Kelly Del Tredici, and Heiko Braak. "A timeline for Parkinson's disease." Parkinsonism & related disorders 16.2 (2010): 79-84.
3. Iakovakis, Dimitris, et al. "Motor impairment estimates via touchscreen typing dynamics towards Parkinson's disease detection from data harvested in-the-wild." Frontiers in ICT 5 (2018): 28.

GET IT ON Google Play

i·PROGNOSIS

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 690494 - i-PROGNOSIS