UNIVERSITY OF BELGRADE

STUDIES AT THE UNIVERSITY
MULTIDISCIPLINARY DOCTORAL STUDIES

Minja N. Belić

Application of artificial intelligence for kinematic signal processing in diagnostics of Parkinson's disease and atypical parkinsonisms

Doctoral Dissertation

Belgrade, 2023

УНИВЕРЗИТЕТ У БЕОГРАДУ

СТУДИЈЕ ПРИ УНИВЕРЗИТЕТУ

МУЛТИДИСЦИПЛИНАРНЕ ДОКТОРСКЕ СТУДИЈЕ

Миња Н. Белић

Примена алгоритама вештачке интелигенције за обраду кинематичких сигнала у дијагностици Паркинсонове болести и атипичних паркинсонизама

докторска дисертација

Београд, 2023

MENTORS AND MEMBERS OF THE EXAMINATION BOARD

Mentors:

Prof dr Zaharije Radivojević, Associate professor, School of Electrical Engineering, University of Belgrade, Serbia

Dr Saša Radovanović, MD, PhD, Principal research fellow, Institute for Medical Research, University of Belgrade, Serbia

Members of the Examination Board:

Dr Milica Đurić Jovičič, Senior research associate, School of Electrical Engineering, University of Belgrade, Serbia

Prof. dr Nataša Dragašević Mišković, Associate professor, School of Medicine, University of Belgrade, Serbia

Prof. dr Milica Janković, Associate professor, School of Electrical Engineering, University of Belgrade, Serbia

Date of defense: 08.11.2023

ПОДАЦИ О МЕНТОРИМА И ЧЛАНОВИМА КОМИСИЈЕ

Ментори:

Проф. др Захарије Радивојевић, ванредни професор, Електротехнички факултет, Универзитет у Београду, Србија

Др Саша Радовановић, научни саветник, Институт за медицинска истраживања, Универзитет у Београду, Србија

Чланови комисије за одбрану докторске дисертације:

Др Милица Ђурић Јовичић, виши научни сарадник, Електротехнички факултет, Универзитет у Београду, Србија

Проф. др Наташа Драгашевић Мишковић, вандредни професор, Медицински факултет, Универзитет у Београду, Србија

Проф. др Милица Јанковић, ванредни професор, Електротехнички факултет, Универзитет у Београду, Србија

Датум одбране: 08.11.2023_____

Acknowledgements

Application of artificial intelligence for kinematic signal processing in diagnostics of Parkinson's disease and atypical parkinsonisms

Abstract: Clinical diagnosis of Parkinson's disease (PD) and atypical parkinsonisms remains a challenging and time-consuming task. This study sought to utilize AI to provide quick support in differential diagnostics, relying on kinematic data obtained from two modalities: repetitive finger tapping and gait.

The main study focus was on finger tapping data obtained by a custom low-weight, low-cost inertial sensor setup. Three groups of patients were recruited, including individuals suffering from PD, Progressive Supranuclear Palsy (PSP), and Multiple System Atrophy (MSA), and a group of healthy controls (HC) without neurological disorders. Statistical analysis of obtained signals showed differences in certain trends between the tested groups, and the utilization of AI models allowed the PD group to be discerned from the controls with accuracy of 92%, whereas all participant groups discerned in a multiclass setting with overall accuracy of 85.18%.

This work also tackled PD diagnostics through the use of AI in analysis of gait, using a sensorized electronic walkway. *De novo* PD patients and a HC group were tested in a series of dual-task tests, where the interference task was of motor or mental type. We were able to programmatically select a subset of gait parameters that best help in PD diagnostics and use the selected parameters to classify PD vs HC group with accuracy of 85%.

Future work should tackle the effect of possible noise in the labels (non-autopsy-confirmed diagnoses), and standardized multi-center data collection that would allow further refinement of the system's predictive power. Recruitment of patients with atypical parkinsonisms for gait-based tests should assess the ability of the proposed analyses to aid in differential diagnostics among these neurological disorders with similar clinical presentations.

Примена алгоритама вештачке интелигенције за обраду кинематичких сигнала у дијагностици Паркинсонове болести и атипичних паркинсонизама

Сажетак: Клиничка дијагностика Паркинсонове болести (ПД) и атипичних паркинсонизама и даље је изазован задатак и изискује доста времена. Циљ ове студије је да кроз употребу вештачке интелигенције понуди брзу потпору у диференцијалној дијагностици, ослањајући се на кинематичке податке прикупљене кроз два модалитета: репетитивно тапкање прстима и ход.

Главни фокус студије су снимци тапкања прстима прибављени помоћу лаганог и јефтиног система инерцијалних сензора. Регрутоване су три групе пацијената, укључујући особе које пате од Паркинсонове болести, прогресивне супрануклеарне парализе (ПСП) и мулти системске атрофије (МСА), као и група здравих контрола без неуролошких обољења. Статистичка анализа прикупљених сигнала је показала разлике у одређеним трендовима у тапкању између тестираних група, а употреба модела вештачке интелигенције омогућила је разликовње ПД пацијената од контрола са ташношћу од 92%, док је све регрутоване групе било могуће класификовати са тачношћу од 85,18%.

Рад се такође бави дијагностиком ПД кроз примену вештачке интелигенције у анализи хода, користећи електронску сензорску стазу. *De novo* ПД пацијенти и група здравих контрола снимани су у серији тестова са двоструким задатком, где је додатни задатак био моторног или менталног типа. Програмски је било могуће одабрати подскуп параметара хода који понајвише доприносе дијагностици ПД, и употребити те параметре за класификацију ПД групе и контрола са тачношћу од 85%.

Будући рад би требало да се бави ефектима евентуалног шума у обележеним дијагнозама јер оне нису потврђене аутопсијом, а прикупљање додатних података стандардизованим протоколом у више клиничких центара омогућило би додатно побољшање предиктивне моћи система. Регрутација пацијената са атипичним паркинсонизмима за тестове базиране на ходу потребна је да се процени способност предложених анализа да помогну у дијагностици поменутих неуролошких обољења.

Кључне речи: Паркинсонова болест, атипични паркинсонизми, кинематичка анализа, вештачка интелигенција, машинско учење, тапкање прстима, анализа хода

Научна област: Биомедицинско инжењерство

Contents

1. INTRODUCTION

Parkinson's disease (PD) is a widespread progressive neurodegenerative disease with a gradual onset, characterized by bradykinesia, rigidity, tremor, and postural instability as major symptoms, the degree of which varies among patients and changes over time [1]. Apart from movement impairment, PD also presents with nonmotor symptoms that comprise neuropsychiatric features and autonomic dysfunction. The disease can be influenced by certain genetic factors, but can also be sporadic, with no obvious inheritance path [2].

Some descriptions reminiscent of PD have been found in traditional Indian texts from 1000 BC, but it was first medically described by a British doctor, James Parkinson in 1817, who called it the "Shaking palsy" [3]. His descriptions read: "Involuntary tremulous motion, with lessened muscular power, in parts not in action and even when supported; with a propensity to bend the trunk forward, and to pass from a walking to a running pace: the senses and intellects being uninjured."

PD is a disease of aging, with a prevalence of 1903 per 100,000 in people older than 80 [4], which is 10 times greater than for people in their 50s. Its burden is expected to grow as the world population ages [5] – the ratio of people over the age of 65 in developed countries is 17.6%, which is a steep climb from 7.7%, the number reported in 1950 [6]. Some reports have already found a 2.5-fold increase in prevalence of PD over the last 30 years [7].

Regardless of the presence of certain identifying characteristics and elaborate clinical criteria, differential clinical diagnosis of Parkinson's disease is still a challenge. When PD clinical diagnoses were contrasted with post-mortem findings, it was found that clinical diagnoses of PD had 88% sensitivity and 68% specificity [8], and when diagnoses were determined in early stages, they were shown to only be correct in 26% of patients who were not treated or were not responding well to treatment. In treatment responsive patients this percentage was 53%. A systematic review paper reported that the overall accuracy of PD clinical diagnosis barely improved in the last 25 years [9], and concluded that new objective biomarkers for *in vivo* diagnosis were urgently required. This high misclassification rate stems from the fact that other conditions have a similar clinical presentation to that of PD, particularly early in the course of the disease.

Getting a diagnosis in early stages could contribute to improved quality of life for the patients by allowing timely introduction of proper therapy and slowing down the progression of the symptoms. Early diagnostics would also be important for clinical trials, allowing patient assignment to the correct group and more robustness in results interpretation. This thesis seeks to contribute to this cause, exploring the role of sensors and algorithms of artificial intelligence in detecting Parkinson's disease in early stages, and assisting differential diagnostics in contrast with similarly presenting neurological disorders.

The thesis will be structured as follows: first an overview will be given in Chapter 1 on Parkinson's disease and atypical parkinsonisms, their hallmarks and differences and similarities in clinical presentations, giving us a peak into the challenges faced in differential diagnostics. Chapter 2 will describe current tests employed in clinical diagnostics, as well as present some approaches using technology, sensors, and algorithms, particularly algorithms of artificial intelligence to try and aid with the diagnostic process. Main hypotheses will be defined in Chapter 3, which we will aim to test in Chapters 4 and 5. Chapter 4 will present efforts on aiding diagnostics through the analysis of kinematic data obtained from the fingers, while Chapter 5 will analyze kinematics of gait. We conclude the work in Chapter 6 and give guidelines for future work.

1.1. Parkinson's disease

Parkinson's disease is reflected in pathological features seen in neuronal tissues, as well as a heap of clinical symptoms, which will be described in this chapter.

1.1.1. Histopathological biomarkers

Histopathological characteristic of Parkinson's disease is a loss of dopaminergic neurons of the substantia nigra projecting to the putamen. These are parts of the basal ganglia, more than half a billion-year-old subcortical brain structure [10] primarily responsible for selection and implementation of purposeful movement, which heavily relies on the release of dopamine. Reduction of dopamine release seen in PD inhibits thalamocortical transmission and results in global movement reduction [11]. The basal ganglia include several interconnected nuclei (Figure 1): the subthalamic nucleus (STN), the striatum (composed of the putamen and caudate nucleus), globus pallidus with internal and external segments, and substantia nigra (consisting of pars compacta – SNc, and pars reticulata - SNr). These segments are densely interconnected, involving direct and indirect pathways, and consisting of excitatory glutamatergic and inhibitory GABAergic projections [12]. These pathways are modulated by dopamine released from the SNc. The SNc is the main source of dopaminergic neurons projecting to other basal ganglia, although there are also dopaminergic neurons spread throughout the striatum. Input from the neocortex gets projected to the thalamic nuclei, which are then projected to the frontal cortex, meaning that the basal ganglia can influence executive functions of the forebrain [13]. Outputs are also sent to the brainstem nuclei involved in motor control. The STN is responsible for receiving input from various brain regions and is thus a good target for deep brain stimulation [14], which was shown to remarkably reduce the motor symptoms in select PD patients. Deep brain stimulation is also delivered to the pallidum, which receives most of the input from the striatum, while the striatum receives the bulk of the input from the cortex. Abnormalities in the basal ganglia pathways lead to loss of movement in parkinsonisms while also being responsible for example for excess movements in Huntington's disease [13].

Apart from their notable role in motion, the basal ganglia also have implications in reward processing, mood regulation and more. Points in this area light up in fMRI scans of people being shown pictures of their object of romantic love [15] and show signs of accelerated aging in those suffering from major depression [16]. Metabolic abnormalities have been observed in basal ganglia in fMRI brain imaging of patients with Tourette syndrome and obsessive-compulsive disorder [13].
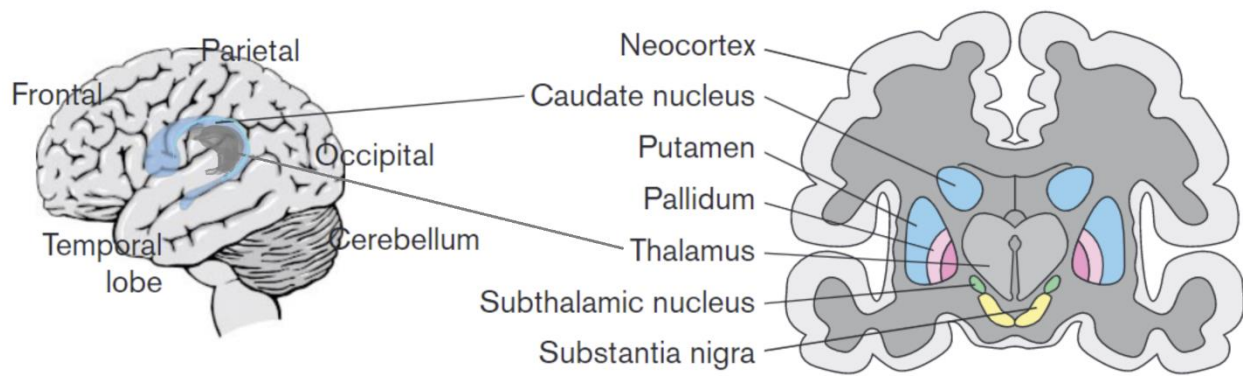
Figure 1 The basal ganglia (Adapted from [13])

Another notable characteristic of parkinsonian disorders is pathologic accumulation of protein tangles within neurons and often glial cells, which happens early in the course of the disease. Neuronal inclusions in PD consist of α-synuclein and are also called Lewy bodies. Lewy bodies are mostly found in SNpc, but also in other brain areas, as well as in the peripheral nervous system, such as the neurons of the enteric plexus of the gastrointestinal tract [17]. A staging system has been proposed for tracking Lewy inclusions from the enteric and autonomic system to the brainstem, and higher parts of the neuraxis [18].

In physiologically normal conditions, α-synuclein is a neuronal protein, widely expressed in the nervous system, and found predominantly in pre-synaptic terminals. Its main role seems to be the control of neurotransmitter release. The misfolded α-synuclein tangles are thought to contribute to PD pathology mainly through aberrant soluble oligomeric conformations – protofibrils – which disrupt cellular homeostasis and affect neuronal death, via different intracellular targets. It also appears to affect its neighboring cells, prompting the spread of the aggregation and progression of the pathology. In a subset of PD cases, a genetic basis of the disease has been identified in a defective SNCA gene coding for α-synuclein [19]. The inheritance of the mutation is autosomal dominant, and generally results in early onset PD. The onset is affected in dose-dependent manner, where the higher number of gene multiplications lead to earlier disease onset.

1.1.2. Clinical features

A common clinical feature of PD is difficulty performing movements, reflected notably in gait, which worsens over the course of the disease. Speed is typically reduced, stride length shortened, double support phase is prolonged, and hesitation is present on start [20]. Walking and standing are complex motor behaviors influenced by mental processes [21], and PD patients are known to have a difficulty performing dual tasks, whether the interference task is motor or cognitive, meaning the problem is not only of motor nature. Inability to deal with simultaneous tasks may be due to central processing becoming too limited, or patients failing to prioritize tasks by importance, not placing balance control over secondary tasks [22]. Basal ganglia, together with multiple other brain areas are thought to contribute to movement automation, serving as the automatic link to other motor areas, connecting sub-movements into automatic skilled motion. This is in line with the observation that voluntary movement is more preserved in PD than automatic movement [23]. Patients can perform normal walking, but they need to direct active attention to it, and external cueing is found to help. This is due to the automatic movement ability being affected, and these attentional strategies allow the motion to rely more on attentional rather than automatic motor control.

Early in the course of PD, when gait alterations may not yet be apparent to the eye, technology can be of help to detect some issues. Spatiotemporal analysis of gait in early-stage PD and healthy controls found that the gait of PD patients is slower, with shorter strides, increased stride duration and double support time, and shorter swing times. Left to right swing asymmetry is also increased and there are inconsistencies in the timing and rhythmicity of gait [24].

Patients with Parkinson's disease experience episodes of sudden and brief inability to produce effective forward stepping, feeling that their feet are glued to the floor. This phenomenon, named freezing of gait (FoG), usually lasts under 30s but some cases may leave patients stuck for several minutes [25]. The occurrence of FoG episodes significantly contributes to falls, which are a frequent event in patients suffering from PD. A study on a cohort of 300 patients with PD in Serbia found that 60% of the patients had had a fall within 6 months prior to testing [26].

Apart from gait, the pervasive motor impairment in PD can be seen as general slowness (bradykinesia) in performing daily activities and delayed reaction times. Patients may have difficulties buttoning their shirts or using utensils; impaired swallowing can lead to drooling [1].

Spontaneous movements such as blinking or arm swinging while walking can be reduced, gesturing can be lost, including loss of facial expressions (hypomimia). This gradually leads to a "mask-like" appearance, which reduces the patients' levels of emotional expressiveness [27] and causes others to misinterpret their emotional state. Embodied simulation theory states that our understanding of the emotions of others is facilitated by facial mimicry. This may also help explain why people with Parkinson's disease have difficulties in recognizing other people's emotions - because of the way the disease affects the contractility of their facial muscles, leading to facial masking [28]. It's been shown that apart from affected expression of emotions, PD patients exhibit impairments in decoding emotional expressions from the faces of others [29]. "The fundamental mechanism that allows us a direct experiential grasp of the mind of others is not conceptual reasoning but direct simulation of the observed events through the mirror mechanism" [30] - says Vittorio Gallese, Italian professor known for the discovery of mirror neurons - neurons that were found to fire in monkeys observing another animal performing an action, just as they fired when the monkey performed the actual action. Indirect evidence points to existence of such neurons in humans too [31].

Abnormalities in handwriting become visible, notably micrographia, where patients write with abnormally small letters. Digital technologies drew the attention to not only micrographia, but also other disturbances in handwriting in PD, including speed, fluency, and acceleration, together known as PD dysgraphia [32].

To assess bradykinesia, clinicians usually use tests which involve rapid, repetitive movements of the hand and foot, such as finger tapping [33], and look for slowness of movement, pauses, and changes over time.

Rigidity is seen in increased resistance to passive movement of a limb, either proximally or distally, and may be accompanied by pain. Rigidity of the neck and trunk may result in abnormal posture and postural deformities, usually late in the disease. Loss of postural reflexes as the disease progresses, together with freezing of gait, contributes to falls and the risk of hip fracture [1].

Tremor in rest is yet another recognizable symptom of PD. It occurs at frequencies between 4 and 6 Hz and hand tremor is of pronation-supination type and resembles the pill-rolling motion [1]. Tremor can be also seen in the legs, lips, chin, but rarely affects the neck and head like essential tremor does.

At the onset of motor symptoms, however, up to 70% of dopaminergic neurons may already be lost. PD often presents with a number of non-motor symptoms first, which may include sleep disorders,

constipation, anxiety, and depression. As the disease progresses, cognitive impairment, dementia, and orthostatic hypotension have been reported [6].

As we will see in the following chapter, PD shares a lot of overlapping symptoms and even some histological features with other movement disorder, which makes differential diagnostics an elusive task.,

1.2. Atypical parkinsonisms

Neurodegenerative disorders in which the degeneration spreads outside of substantia nigra and is more extensive than in Parkinson's disease are categorized as atypical parkinsonisms [2]. Their manifestations resemble PD, but are usually more complex. As in PD, Lewy pathology is also seen in Multiple System Atrophy (MSA), while Progressive Supranuclear Palsy (PSP) is considered a tauopathy, presenting with tangles of tau protein, alongside Cortico-Basal Degeneration, Guam Parkinson-Dementia Complex and Chronic Traumatic Encephalopathy. A category of parkinsonisms associated with TDP-43 proteinopathy has also been recognized [34].

1.2.1. Progressive supranuclear palsy

Progressive supranuclear palsy (PSP) is a disorder of tau protein aggregation, first described in 1964 by Steele, Richardson and Olszewski [35], prior to which it was commonly misdiagnosed as Parkinson's disease. PSP is assessed to be present in 5 to 6 out of 100 000 people [36], making it much less common than PD. It is a progressive neurological disease notable for supranuclear ophthalmoplegia mainly in vertical gaze, pseudobulbar palsy, dysarthria, and dysphagia that leads to aspiration, dystonic rigidity of the neck and upper trunk, severe gait disturbances and balance impairment, but also less commonly or less constantly sleep disturbances, depression, urinary incontinence, constipation, apraxia, tremor, and more [36].

Particularly in earlier stages, the characteristics of PSP, especially the parkinsonian subtype, resemble those of idiopathic PD [37]. It is common that this diagnosis is given after 3 to 4 years from onset, once the typical manifestations of the condition - supranuclear gaze palsy and falling – come to light [38]. PSP patients have a poorer response to levodopa, with this treatment being effective in only 26% of cases [39], and these differences in medication response have been themselves used to aid diagnostics, although with limited success [40].

1.2.2. Multiple system atrophy

Another condition with some overlapping symptoms with PD is Multiple System Atrophy (MSA). When 134 patients with clinically diagnosed MSA were analyzed, it was found that for 83 of them (62%) the diagnosis was confirmed post-mortem [41]. For the remaining subjects, the disease was mostly confused with dementia with Lewy bodies, followed by PSP and PD.

MSA is a rare parkinsonism, involving pathologic accumulation of aggregated α-synuclein as glial cytoplasmic inclusions [42]. Main features of this disease are parkinsonism or cerebellar ataxia with autonomic failure. Two phenotypes are recognized: MSA with predominant parkinsonism (MSA-P) and MSA with predominant cerebellar ataxia (MSA-C). Other clinical signs include early postural instability, dysphagia, orthostatic hypotension, and urinary incontinence [42]. There are other conditions that may sometimes be confused with PD, including dementia with Lewy bodies, Alzheimer's disease, postencephalitic parkinsonism and essential tremor [43], but those disorders are out of scope of this work.

2. DIAGNOSTICS OF PARKINSON'S DISEASE AND ATYPICAL PARKINSONISMS

PD and atypical parkinsonisms, as well as some other neurological disorders, manifest often with an overlapping palette of symptoms. Frequency of occurrence of these symptoms can be and is used in clinical practice for differential diagnostics: e.g. asymmetry in motor signs is almost always present in PD, but only sometimes in MSA and PSP. Autonomic disfunction is always present in MSA, but only sometimes in PD and almost never in PSP. Dementia is almost always seen in PD and PSP, but rarely in MSA. Treatment with levodopa can also be used as a differentiating marker, as PD patients almost always respond to treatment, while PSP and MSA patients are less responsive [34]. Some of these features are represented as clinical diagnostic criteria in the UK PD Brain Bank criteria [43]. Clinimetric rating scales have been developed for staging and progress assessment of parkinsonian disorders, the most prevalent of which are the UPDRS and Hoehn & Yahr scales, described briefly in Chapter 2.1. In Chapter 2.2. we give an overview of approaches utilizing technology and artificial intelligence to aid in differential diagnostics of PD.

2.1. Clinical motor impairment, disease staging and diagnostic tests

UK Parkinson's Disease Society Brain Bank lists PD diagnostic criteria in three steps. Step one includes bradykinesia, and one or more of the following: muscular rigidity, 4-6Hz rest tremor, postural instability not associated with other known dysfunctions of the visual, vestibular, cerebellar, or proprioceptive system [43]. Step two lists the exclusion criteria for the diagnosis of PD, and these include a history of repeated strokes, head injury and definite encephalitis, oculogyric crises, use of neuroleptic medications at presentation, more than one relative affected, sustained remission, features remaining unilateral after 3 years, supranuclear gaze palsy, cerebellar signs, early autonomic involvement, severe dementia, Babinski sign, cerebral tumors, negative response to high doses of levodopa, MPTP exposure. Step three involves supportive criteria for PD diagnosis, where three or more are required to confirm PD: unilateral onset, progression, persistent asymmetry, very responsive to levodopa, responsive for at least 5 years, clinical course of 10+ years, and severe levodopa-induced chorea.

Unified Parkinson's Disease Rating Scale (UPDRS) is the most widely used clinical rating scale for PD. It was originally developed in 1980s and revised in 2008 [44] by the International Parkinson and Movement Disorder Society (MDS) [45]. The scale includes 4 parts: Part I - Non-motor Aspects of Experiences of Daily Living; Part II – Motor Aspects of Experiences of Daily Living; Part III – Motor Examination; Part IV – Motor Complications;

Part I is a questionnaire assesses the patient's complex behavior, including cognitive impairment, hallucinations and psychosis, depressed or anxious mood, apathy, features of dopamine dysregulation syndrome (detected through the presence of unusually strong and hard-to-control urges, e.g. towards gambling), sleep problems, daytime sleepiness, pains and aches, urinary and constipation problems, light headaches on standing, fatigue. All items are scored by intensity and frequency of occurrence on the scale from 0 (none) to 4 (severe).

Part II is a questionnaire that examines speech problems, drooling, chewing and swallowing, troubles eating (such as weakness while holding utensils) or dressing, hygiene, handwriting, engaging in hobbies and other activities, turning in bed, presence of tremor, getting out of bed or a chair, walking and balance, freezing of gait.

Part III is a motor examination administered by a specialist of neurology. The patient is asked to perform a number of motor tasks and the examiner rates what they see on a scale from 0 (normal), through 1 (slight), 2 (mild), 3 (moderate), to 4 (severe). This part involves speech examination, where the examiner listens to the patient freely talking, and pays attention to the volume, prosody and clarity, slurring, palilalia (repetition of syllables), and tachyphemia (fast speech, squashing syllables together). The next on the examination list are facial expressions, both while talking and while being quiet, looking for signs of masking. Rigidity is judged by slowly moving the patient's neck and limbs, while they are instructed to keep limp.

Finger tapping is the fourth test in the motor part of the scale. The patient is instructed to tap their index finger against the thumb repeatedly as quickly and as widely as they can. Motor performance degrades much more easily during sequential motions in case of individual finger opposition than non-individual finger oppositions, and it is therefore standard to perform this test with index-thumb oppositions [46]. Both hands are assessed, and attention is paid to speed, amplitude, hesitations or halts and decrements in amplitude. Hand movements are tested on both hands separately, by having the patient clench and open their fist repeatedly, and watching the speed, amplitude, hesitations or halts, and decrements in amplitude. Pronation and supination of hands are also tested by having the patient extend their arms, and repeatedly turn the palm upwards and downwards alternately, trying to achieve maximal speed and extent. Finger tapping has been shown to better correlate with the overall UPDRS scores than hand pronation and supination or hand opening and closing [46].

Tapping of toes is tested by having the patient sit with their feet on the floor, and then tap their toes on each foot separately as fast and with as large an amplitude as they can, evaluating speed, amplitude, hesitations or halts and decrements in amplitude.

Agility of the legs is tested on each foot separately, by having the patient sit with their feet on the floor and then repeatedly raise and stomp. Arising from the chair is tested next, by having the patients stand up from the chair while having their arms crossed on their chest, observing difficulties while doing so, as well as posture after standing up. Gait and freezing of gait are tested by having the patient walk back and forth from the examiner for at least 30m, assessing stride length, speed, level of foot lifting, heel strike, turning, arm swinging, and presence of hesitation, stuttering, and freezing of gait – particularly at turning points and end of the task. Postural stability is tested by observing the response to abrupt pull on the shoulders in standing position, and posture is assessed during standing, getting up from a chair and walking. Body bradykinesia rating is given based on slow, hesitant movement and lack of movement in general during spontaneous activity. Postural hand tremor is rated for each hand separately, by watching the patient stretch out their hands in front of them for 10s, palms facing down. Kinetic hand tremor is tested by having the patient touch their nose and then the examiner's finger. The patient is observed for rest tremor during the exam, for all limbs separately, and lips and jaw. Consistency of rest tremor is also rated. Presence of dyskinesia, such as chorea and dystonia are noted. This part also includes noting the **Hoehn and Yahr stage** [47], which is a scale that defines broad categories of motor function. The stages are defined from 0 to 5, given as follows:

0 – Asymptomatic

1 – Unilateral involvement

2 – Bilateral involvement without balance impairment

3 – Mild to moderate bilateral involvement; some postural instability; physically independent; assistance required to recover from the pull test

4 – severe disability; still able to walk or stand unassisted

5 – wheelchair bound or bedridden unless aided

Modified Hoehn and Yahr scale [48] adds also 1.5 – Unilateral and axial involvement and 2.5 – Mild bilateral disease with recovery on pull test.

Part IV of the MDS-UPDRS scale deals with motor complications, based on historical and objective information, assessing dyskinesias and their prevalence during the patient's waking day, and their functional impact on daily activities. Time spent in the OFF state (on PD treatment but experiencing some hours of slowness, shaking or similar) is rated as well, from no off time to more than 75% of the day spent in off time. The scale rates the impact these motor fluctuations on the patient's daily functioning, and their predictability. For patients with motor fluctuations, the scale rates the percentage of OFF time that includes painful dystonia.

2.2. Technology and artificial intelligence in aid of diagnostics

There have been numerous approaches to employing technology to aid in detecting and differentiating PD [39], perhaps most famously neuroimaging. Applications of kinematic analyses of gait or hand movements have been increasing, both for diagnostic support and for disease stage monitoring. Other sensing modalities, such as EEG or facial analysis, have also found a use case in assisting PD diagnostics.

Artificial intelligence (AI) and machine learning (ML) algorithms have been vastly relied on to build upon the various sensors and help offer diagnostic predictions on an individual level. ML algorithms have been shown to give a significant contribution to improvements in classification between patients with PD and people without neurological disorders, compared to results obtained in the clinical practice alone, although published studies involved varying numbers of participants, stages of the disease, particular analysis and instrumentation used, and demonstrate better results when the PD participants are in more advanced stages of the disease. Larger numbers of recruited participants also lead to better results.

2.2.1. Neuroimaging

Many studies have turned to neuroimaging methods to find group level differences between neurodegenerative conditions [49], and have deepened our understanding of the pathophysiology. Distinguishing macroscopic features in neuroimaging have been discovered between MSA, PD and PSP, such as atrophy of the pontine base in MSA, but not in PD and PSP, or marked atrophy in the superior cerebellar peduncle which is not present in the other two disorders [34]. MRI is a frequently used imaging technique, particularly the t1-weighted modality [50]–[58], although computer aided diagnostics relying on diffusion tensor imaging (DTI) [59], [60] and susceptibility weighted imaging (SWI) [61] can also be found in literature[1]. Some effort has been put into leveraging multimodal data for more reliable computer aided diagnostics, combining t1-weighted, t2-weighted and DTI [62]–[65][62]–[65]. While t1-weighted imaging has had some trouble discerning PD from healthy controls [51], [54], as the brains of those with early and mid-stage PD usually appear normal, Planetta et al [60] achieve perfect separation using DTI, although having also included clinical data.

To improve the predictive power of neuroimaging methods when it comes to the clinical outcome of individuals, machine learning algorithms coupled with large neuroimaging databases have achieved some quite exciting results, reaching accuracies over 90% [50], [58], [59], [66]–[71], and discerning PD from PSP and MSA [54], [60], [63]–[65].

The most widely used classification algorithm is the support vector machine (SVM) (More on SVMs can be found in Chapter 4.5.1). SVM papers report very good, or even perfect results for binary classifiers, but sadly multiclass classification does not match those numbers. Gong et al [72] believe that the road to better SVM performance is through optimized kernel selection. They expand on the large margin distribution machine (LDM) algorithm, introduced by Zhang & Zhou [73], based on the notion that not

---

[1] MRI (magnetic resonance imaging) is an imaging technique that uses a powerful magnetic field to polarize protons in scanned tissues parallel to the field, and a transverse radiofrequency pulse that pushes the protons out of the equilibrium. To construct an image, T1-weighted imaging uses the tissue-specific time it takes for protons to realign with the magnetic field once the RF field is off (longitudinal relaxation time), while T2-weighted imaging relies on the duration of proton precession (transverse relaxation time) in response to the RF signal. DTI measures diffusion of water molecules in tissues in various directions, and SWI exploits substance differences in magnetic susceptibility.

only is maximizing the minimal margin important, but also taking into consideration the distribution of the margin – its mean and variance. They employ a deep neural network for kernel mapping, pre-training it with a Restricted Boltzmann Machine (RBM) and then fine tuning in a supervised manner, to finally employ an LDM for classification. They illustrate the efficiency of this approach on the problem of discerning PD from controls on two datasets: MRI and transcranial sonography (TCS), in both cases outperforming classical SVM and LDM approaches.

Others try to improve SVM classification accuracy through selection or extraction of different features. The go-to are often volumetric measures derived from regions of interest (ROI) [56], incorporating thus a-priori knowledge on the manifestations of the disease. The input dimensions are sometimes reduced through PCA [50], [52], [58]. Morisi et al [65] include features generated through a graph-based method. Similarly, Peng et al [57] use multilevel features – low level volume and thickness, and high-level correlative features. Amoroso et al [70] use connectivity features and further filter them through a random forest. Adeli et al [67] focus on selecting features that best benefit the classification scheme in the kernel space, unlike conventional techniques that select features based on their performance in the original input feature space. They have, however, combined MRI with SPECT (Single Photon Emission Tomography) data, and SPECT turned out to be the main carrier of discriminative information. Including clinical data has been demonstrated to improve results [60], [70], and interestingly, so has addition of information on gender and age [55].

The issue with imaging analyses is that they require a lot of time and money, and the modalities that offer the highest reliability expose the patient to ionizing radiation and cannot be used in people who are hypersensitive to the radioactive dye. While MRI has the advantage of not (necessarily) using radioactive substances, dopaminergic images obtained by SPECT can detect Parkinson's Disease at an early stage and have demonstrated superior classification ability in discerning PD from healthy subjects. When SPECT images were employed in combination with deep learning methods, concretely a convolutional neural network (CNN), a stunning 100% accuracy was achieved by Esmaeilzadeh et al [71] on validation and test sets.

This and other studies using deep learning algorithms [66], [68], [69] have been greatly facilitated by the mass collection of data, basing their research on medical images obtained across continents by the Parkinson's Progression Markers Initiative (PPMI) [74]. PPMI is a large-scale, international public study with the aim of identifying PD progression biomarkers. They make SPECT and MRI recordings, together with clinical and biological data, available to the research community, on condition they provide reports on their findings. The work by Kim et al [69] is interesting because it shows that it is possible to have a limited set of SPECT images and still come to decently accurate PD vs controls classification, through transfer learning from an abundant set of non-medical images.

2.2.2. Kinematic analysis

Although neuroimaging offers superior results in terms of diagnosing PD, its downsides such as higher time consumption, price and limited availability in less urban areas, as well as exposure to ionizing radiation in some modalities, have prompted the development of alternative technological approaches to diagnosing and monitoring PD and other movement disorders, in particular the use of various sensors to track movement of the body or its parts. Kinematic analysis has proven to be useful in diagnostic aid and motor function assessment, through recording and analyzing motion from upper or lower extremities, or

a multitude of sensors capturing whole body motions. In upper extremities, particularly bradykinesia and hand tremor have been analyzed [75], [76]. Diagnosis has relied on hand motions, captured by computer keyboards [77]–[79], wearable inertial sensors [80] or sensors integrated in smartphones [81]–[83], electromagnetic sensors [84], and cameras [85], by which researchers sought to build upon traditional tests performed in clinical settings for diagnosis of PD. This primarily encompasses finger tapping and tremor tasks (action/rest) that are a part of the standardized UPDRS battery of tests, part III [86], as described in Chapter 2.1. Although UPDRS is an indispensable tool for rating PD, it is still prone to examiner subjectivity and inter-rater variability. Motor dynamics in repetitive finger tapping have thus been quantified in research through the utilization of sensors such as accelerometers and gyroscopes, to try and make the motor progression tracking more objective [46]. Concerning lower extremities, analysis has relied on data collected using force sensors inside the shoes or integrated into a walkway [87]–[89]. Electronic walkways, such as GAITRite, although expensive and tied to a laboratory setting, have proved to be reliable and have often been used to analyze gait in parkinsonisms, due to a large number of sensors they contain and abundance of spatio-temporal variables they offer, as well as the possibility of analyzing individual footsteps [90]. Infrared cameras, Kinect or motion capture with markers have also been used [91]–[94], as well as inertial sensors placed on the legs, waist or feet or smartphone integrated sensors [95]–[101]. As inertial sensors are of particular importance for this thesis, let us dedicate some extra attention to them. Inertial sensors are a widely adapted means of recording movement of an object in an inertial reference frame. Though they may come as single sensors, they are usually packed into an inertial system, an inertial motion unit (IMU) that contains an accelerometer, a gyroscope and potentially a magnetometer, each measuring motion on 3 axes. They could be manufactured in different technologies, such as mechanical, quartz or MEMS (Micro electromechanical systems) technology, spanning a large range of prices, with the navigation grade sensors used for military purposes reaching prices of over $100 000, while consumer IMUs, such as those built into smartphones, tablets or gaming systems can cost less than $10 [102]. The recent steep reduction in price driven by highly integrated MEMS technologies, prompted wide adoption of inertial sensors in various consumer products, but also health monitoring applications.

The basis of MEMS systems is a proof mass suspended on a spring km, as shown In Figure 2.
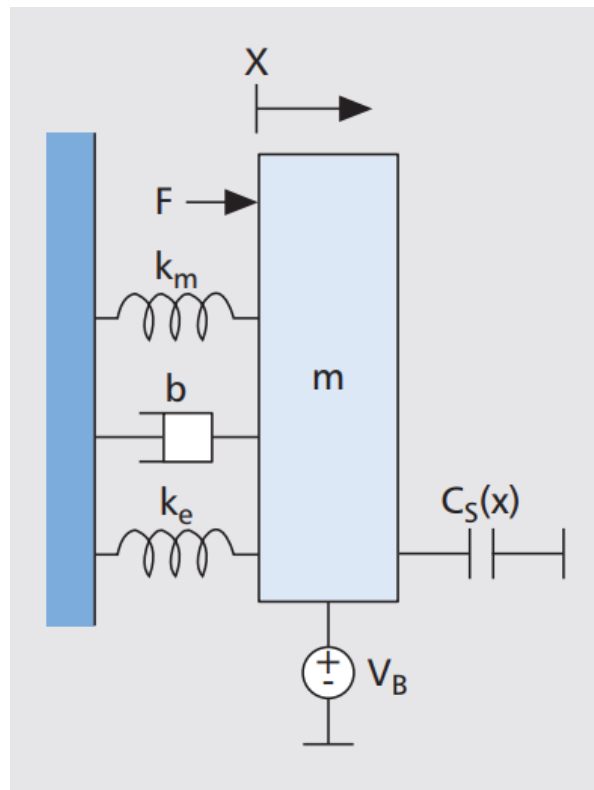
Figure 2 Electromechanical spring/mass system in MEMS sensors (Image adapted from [103]*)*

The input force acting on the mass, i.e. the quantity to be measured, displaces the mass, and this displacement is used as a proxy to measure the input force. This displacement can be measured using a variable capacitor Cs(x), one side of which is attached to the proof mass, and the other is fixed. A change in the position of the mass causes a change in charge as described by Equation 1 [103], where Vb is a fixed bias voltage.

$$\Delta Q(\Delta x) = \frac{dC_S(x)}{dx} V_B \Delta \tag{1}$$

Accelerometers use this sort of simple spring/mass system to measure linear acceleration, typically given in gees, where 1g = 9.81m/s2, representing acceleration due to gravity of the Earth. The input force to be measured is a result of gravity (which is why accelerometers are used in mobile phones for sceen orientation) or linear acceleration. Gyroscopes, on the other hand, require a more complex system (Figure 3). Gyroscopes measure angular velocity in degrees/second or radians/second. They require the proof mass to be vibrating, and detect the Coriolis acceleration proportional to rotational velocity along an axis orthogonal on the axis of vibration. The amplitude of this oscillation must be regulated so that the speed is maintained stable and known. The system has two springs and has two degrees of freedom – sensing (x) and drive(y). A constant speed of rotation results in a proportional amplitude of periodic displacement of the proof mass, phase shifted by 90 degrees from the drive position, as the Coriolis acceleration is proportional to its derivative, the drive velocity. This produces an amplitude-modulated flow of charge in the sensing capacitor.
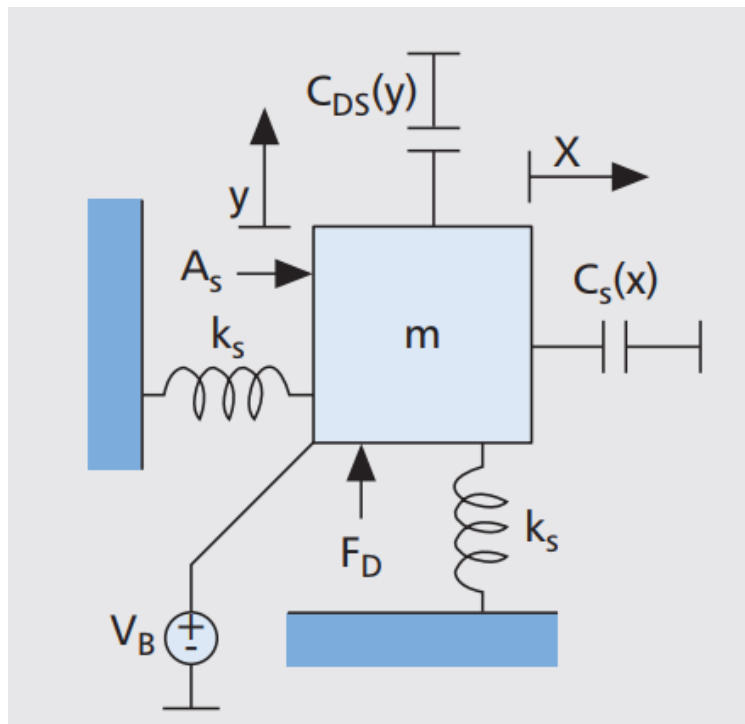
Figure 3 A MEMS gyroscope schematic (Image adapted from [103]*)*

Similar hardware components are nowadays embedded in smartphones and other smart devices, opening a window to the utilization of already available hardware for aid in diagnosis and progression monitoring in PD and other disorders. In combinations with signal processing and custom applications, smartphone built-in sensors have been used, although somewhat awkwardly, to quantify certain PD symptoms, e.g. bradykinesia [104].

Technology-assisted diagnostics and disease monitoring based on kinematic data obtained from upper extremities focus particularly on bradykinesia and hand tremor, the assessment of which may include motion tracking from video captures or inertial motion units, but can also rely on smartphones and the analysis of typing on a virtual or physical keyboard[2]. Machine learning approaches have been of help in these analyses, as illustrated in Figure 4, with the upper panel displaying the success of ML applications in diagnostics, and the lower panel pertaining to disease monitoring applications. A specific marker is given to each study, where instrumentation used is coded in marker shape, number of patients in marker size, and color representing reported performance metrics (sensitivity – Se, specificity – Sp, and accuracy - Ac). The figure presents the most frequently used or the most successful ML algorithms based on reported performance, rather than all of them. The columns show different ML algorithms, sorted in the alphabetical order: ANN – Artificial Neural Network, ENS – Ensemble of different classifiers, EVOL – Evolutionary algorithm, kNN – k Nearest Neighbours, LR – Logistic Regression, NB – Naïve Bayes, SVM – Support Vector Machine and TREE – Tree based algorithms.

If we look at the sensitivity and specificity of the algorithms used for diagnostic aid using smartphones, we see that the ratio of Se/Sp for NB, SVM and LR are respectively: 56%/100%, 56%/100% and 74/100%, meaning that high confidence can be put into identifying healthy participants, but detections of persons with the disease are often missed [83]. BAG DT, AdaBoost and C4.5 approaches had a mutually similar performance, showing better sensitivity, but lower specificity compared to the previously mentioned algorithms, with Se/Sp ratios of 82%/90%, 83%/85% and 83%/75% respectively. Accuracy of 95% was reported for early diagnostics based on tremor analysis via ANN and a smartphone [81]. Sensitivity and specificity of 96% and 97% was reported using an ensemble of eight different ML algorithms used on top of keyboard typing data [78]. Giancardo et al. [77] and Arroyo-Gallego et al [79], who also used a keyboard, but opted for SVR algorithms for early diagnostics, report Se/Sp ratio of 71%/84% and 77%/72% respectively. It should be noted that these measurements were taken by a standard smartphone or regular PC keyboard, suggesting that implementation of test procedures is possible with globally available and affordable equipment.

---

[2] A survey of AI applications in PD diagnostics has been published in [105]
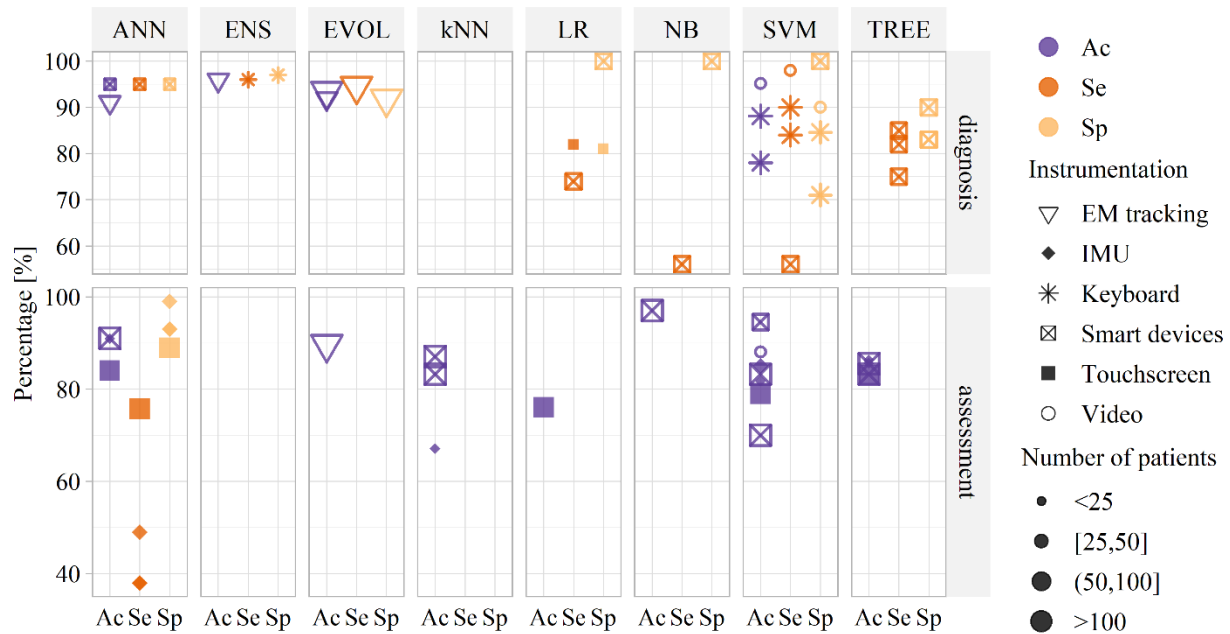
Figure 4 Performance measures of different ML algorithms applied on kinematics of upper extremities [Published in [105]*]*

Intelligent algorithms have also been used to assess and track symptoms of motor impairment, with the purpose of providing a more objective and automated means of assessment of symptom severity than what is achieved by tests in the clinical settings alone. Researchers have used tasks such as those used in clinical testing: tremor, finger tapping, or other repetitive movements, and extracted features such as tremor frequency [106] or amplitude and frequency of repetitive motions [104]. The results of these algorithms are presented in Figure 4, in the lower panel, and include the most used supervised machine learning algorithms based on upper body measurements in PD.

Stamate et al [107] have developed a mobile application, CloudUPDRS, which helps in recording UPDRS motor tasks, and employs a deep learning model to discern between high quality and low-quality recordings. Machine learning has also been deployed in applications that aim to extract useful information from everyday activities of PD patients and use them to provide meaningful insight into how well a patient is responding to therapy and what the next course of action should be [108], [109]. Fisher et al [109] used ANN on data from an accelerometer mounted on the wrist to automatically detect sleeping states, ON and OFF states and dyskinesia in home settings. They found that ON/OFF states and dyskinesia can be detected with sensitivity and specificity above 80%. Hammerla et al [108] achieve comparable results using deep learning. Detection and classification of dyskinesia and bradykinesia have been tackled using both traditional machine learning [110], [111] and deep learning [111] to achieve high accuracy (84 to 90%).

Analysis of gait has been the staple of studying, monitoring and diagnosing PD. Some studies have relied on infrared motion capture systems, others on sensorized electronic walkways. Inertial sensors have also emerged as a cheaper and lighter alternative, with a varied number of sensor units mounted on the legs, with effort put towards simplification of the setup and sensor number reduction [112]. Artificial intelligence has often been used in development of systems for diagnostic and monitoring aid based on the data describing motion of lower extremities [75]. A study that focused on early PD diagnostics and used a kNN classifier [101] achieved accuracy of 85.5%. Using SVM also reached 85% accuracy or higher, as shown by a couple of studies [88], [113]. LDA (Linear discriminant analysis) used in this context provided Se/Sp of 88%/86% for diagnosing early PD, where early was considered to have a UPDRS score below 15, and this was done using gait data from three different tests [97]. Sensitivity and specificity increased to 100% in the same study when considering UPDRS scores above 20. Se, Sp and Ac achieved by an RBF NN applied on a PhysioBank dataset of 93 PD patients in the early to moderate stage, were 96.77%, 95.89% and 96.39% respectively [89]. Most used supervised machine learning algorithms for diagnostics on lower extremity data are shown in the upper panel of Figure 5, while the lower panel displays results obtained for assessment using kinematic data. The results are presented similarly to Fig 4.
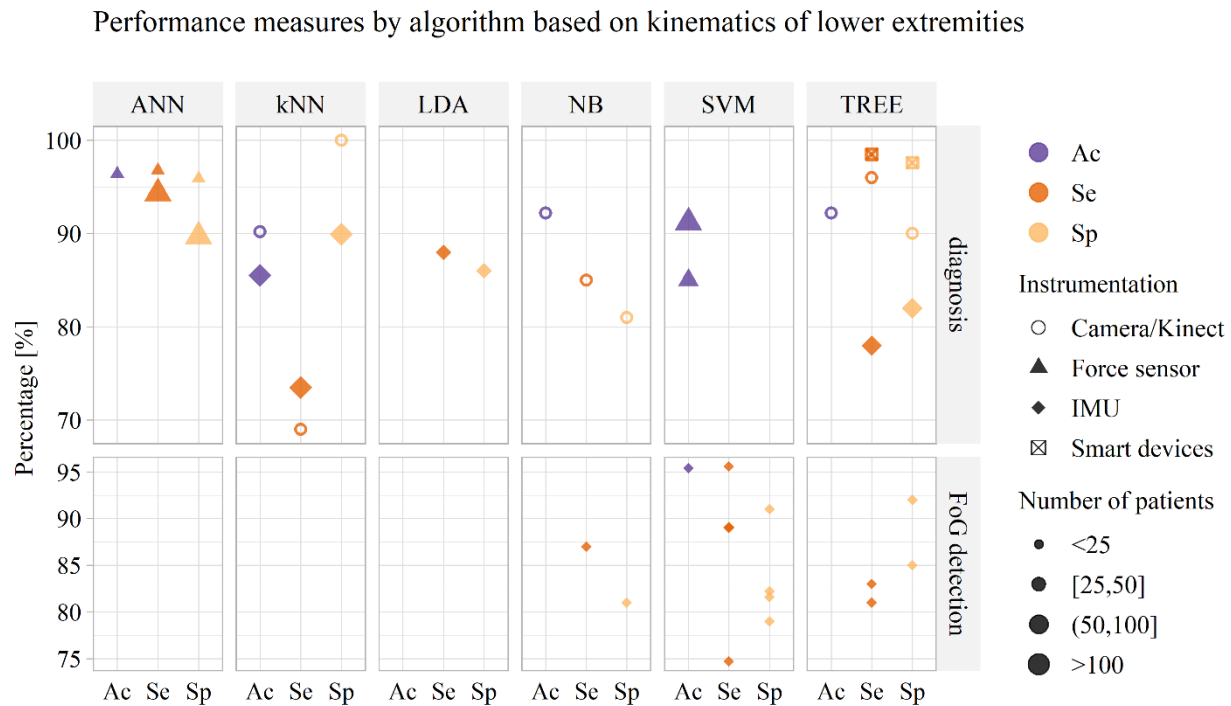
Figure 5 Performance measures of different ML algorithms applied on kinematics of lower extremities [Published in [105]]

A major niche for applications of artificial intelligence is freezing of gait (FoG) and falls, and wearable sensors have found their way to detecting and predicting these states [114], due to their ability to be used for home monitoring, where they can track gait during walking along complex paths with a lot of stops and turns. Commonly, the sensors are attached on the legs and waist [115]–[117], with some studies including wrist and chest wearables [118], [119]. Smartphones [120], with their integrated IMUs, and

smartwatches [121], have also been used as a means of unobtrusive FoG detection and gait-based diagnostic aid.

To detect FoG, scientists have often employed the SVM algorithm. SVM with a polynomial kernel achieved Se/Sp of 89%/91% for FoG detection, and 75%/88% for FoG prediction [117]. The same study tested other ML algorithms as well, including ANN, NB, RF, kNN, linear kernel SVM and Extreme Gradient Boosting, but the Polynomial SVM outperformed them. In another paper [116], the researchers designed personal and generalized FoG detection models, using SVM, and reported that the personalized model with Se/Sp of 88.1%/ 80.1 % performed better than the generalized one with Se/Sp of 74.7%/ 79%. Introducing nonlinear SVM kernels slightly improved FoG detection accuracy compared to the linear SVM (Ac=95.4% compared to 94.2%) [115]. In order to predict incoming FoG episodes, Mazilu et al [122]used unsupervised learning to identify patterns in kinematic data and derive features for predictions. The lower panel of Fig 5 summarizes the results for the most used machine learning algorithms on gait data for FoG detection.

The problem of frequent falls in PD has been addressed using smart wearable devices and camera-based systems, and here machine learning algorithms used for this purpose have shown to have better prediction rates than threshold-based approaches [123].

Jane et al [124] automated detection of severity of gait impairment, in accordance with the H&Y scale, with the help of wearable sensors and a Q-backpropagated time-delay neural net, achieving accuracy of about 90%. Se and Sp of over 90% were also achieved using an SVM classifier on accelerometer data [125]. Machine learning was also employed to assess the effects of deep brain stimulation (DBS) on ground reaction force data, specifically LR, SVM and PNN (Probabilistic Neural Network), and a positive effect of DBS was shown on walking patterns [126].

Motion analysis using data collected from upper and lower extremities combined has been performed with the aim of helping PD diagnostics [92], [96], [99], staging [99], [100], detection and assessment of dyskinesia [127], bradykinesia [128], [129], tremor [127], [129], [130], and FoG [118].

A combination of sensors has been used to collect data from everyday activities to try and discern healthy controls from patients with PD, either with cognitive impairments or without them, and a group of patients suffering from cognitive impairment due to disorders other than PD [92]. It was shown that these four categories can be classified with 86% accuracy using AdaBoost DT algorithm. Comparing only the healthy controls with PD patients, using inertial sensors mounted on the upper and lower body, separation accuracy ranges between 79.62% and 84.1% for NB, kNN, LDA, SVM and DT, with an ensemble of classifiers surpassing 90% [100]. When SVM was employed for healthy vs PD classification, splitting patients into groups with H&Y stage I, II or III, achieved accuracy was 94.5%, 87.75%, 93.63% respectively. In other studies, EML (Extreme machine learning), PNN and kNN were used for diagnostics [99], and HMM (Hidden Markov model) for estimation of tremor severity based on upper and lower limbs, with an accuracy of 87% [130]. SVM[129] and DNN (Dynamic neural network) [127] helped to assess bradykinesia, tremor, and dyskinesia with over 90% certainty. Similar results were achieved for FoG detection with NB, RF, DT and RT classifiers. Figure 6 presents the most often used or most successful ML approaches used to assess and diagnose PD using a combination of sensors on the upper and lower extremities.

Performance measures by algorithm based on combined upper and lower extremity kinematics
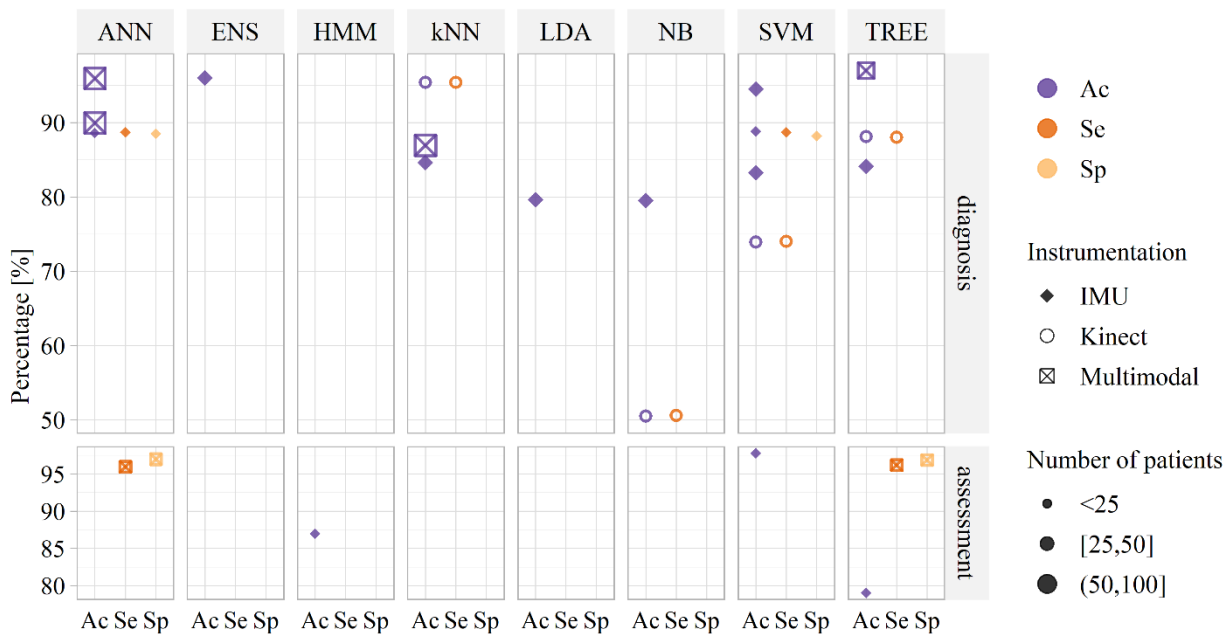
Figure 6 Performance measures of different ML algorithms applied on kinematics of combined upper and lower extremities [Published in [105]]

Table 1shows a summary of papers that achieved promising results using machine learning for diagnosis and assessment of patients with PD. Using different systems gave comparable results, although cameras and motion capture systems have a limited use, due to their price as well as requirements for a dedicated recording space, while wearable sensors have a much lower price range and can be used for monitoring

in various environmental conditions. Smartphones with their integrated sensors are also used often since they are globally available and relatively affordable and have an integrated capability of wireless data transmission [131]. Wearable devices appear to have a large potential for use in telemonitoring of PD. We've seen machine learning algorithms employed to help in two main streams: offering diagnostic aid and disease assessment and monitoring. The algorithms commonly used are SVM, kNN, NB, ANN, LDA, DT, RF, although others have been used too. A notable number of studies have shown that simple wearable sensors combined with machine learning algorithms can form a powerful support tool for PD diagnostics and assessment. Neural networks have been used less often, possibly due to them being data-greedy, and medical data is not always easy to come by, so some studies used simulated data (e.g. falls) [123]. A potential issue with wearable sensors and machine learning is the need for data annotation in home environment scenarios, and potential errors in data labelling based on clinical evaluation in early stages, limiting the use of strictly supervised machine learning approaches.

Table 1 A selection of papers providing promising results for different applications based on movements of different body parts, using different instrumentation, protocols, and algorithms.

Ref.	Goal	Type of observed motion	Body part	Instrumentation	Subjects	Algorithm	Best performance [%]		
							Sp	Se	Ac
[132]	Diagnosis	Finger tapping	Up	EM tracking	107 PD, 49 HC	EVOL	91.8	94.6	93.5
[78]	Diagnosis	Typing	Up	Keyboard	20 PD (mild), 33 HC	ENS	97	96	
[81]	Diagnosis	Arm movements at rest, waving and walking	Up	Smartphone	21 PD (>1 year), 21 HC	ANN	95	95	95
[85]	UPDRS scoring	FT	Up	Video	13 PD (UPDRS: 0-3)	SVM			88
[106]	UPDRS scoring	Hand tremor	Up	Smartphone	52 PD	NB			97
[132]	UPDRS scoring	FT	Up	EM tracking	107 PD, 49 HC	EVOL			≥89.7
[89]	Diagnosis	Gait	Low	Force sensor	93 PD (mild and early), 73 HC	ANN	95.9	96.8	96.38
[95]	Diagnosis	Gait, Posture	Low	Smartphone	10 PD, 10 HC	RF	97.6	98.5	98.0
[101]	Diagnosis	Gait	Low	IMU	156 PD, 424 HC	kNN			85.51
[115]	FoG detection	Gait	Low	IMU	20 PD (H&Y>2)	Linear SVM	95.6	82.2	95.4

[94]	Diagnosis	Gait	Low	Camera system & Force plate	23 PD (H&Y: 2), 26 HC	RF	90	96	92.6
[52]	Classification of severity of motor disorders	Unconstrained activity	All	Multimodal	19 PD, 4 non-PD	ANN	97.1	94.9	
[129]	Assessment	FtN, FT, HOC, HT, SIT, HA	All	IMU	12 PD (H&Y: 2-3)	SVM			>95
[128]	Diagnosis	Gait, Posture, FT, RT	All	Smartphone	10 PD, 10 HC	RF	96.9	96.2	
[100]	Diagnosis (PD - H&Y I)	Gait	All	IMU	27 PD (H&Y:1-3), 27 HC	SVM			94.5

PD – Parkinson's disease; HC – Healthy controls; UPDRS – Unified Parkinson's disease Rating Scale; H&Y – Hoen and Yahr scale; ANN – Artificial Neural Network; EVOL - Evolutinary; ENS – Ensemble; kNN – k-Nearest Neighbours; NB – Naïve Bayes; SVM – Support Vector Machine; RF – Random Forest; Ac – Accuracy; Se – Sensitivity; Sp – Specificity; IMU – Inertial Measurement Unit; EM – Electromagnetic tracking; FtN – Finger to nose; FT – Finger tapping; HOC – Hand opening/closing; HT – Heel tapping; SIT – Sitting; HA – Hand alternating; RT – Reaction time.

2.2.3. Other modalities

Apart from a large pool of research papers presenting the use of neuroimaging and kinematic analyses to help manage and understand parkinsonisms, a number of studies have attempted to utilize technology to capture other features of PD, among which is the use of EEG to detect and potentially predict episodes of freezing of gait, and visual analysis of the face to quantify emotional expressiveness. This list of various technological approaches in PD is not comprehensive [133], but going wider into this topic would take away from the main focus of this work. On the other hand, it should be noted that the focus of this work is on kinematics, and the overview of other modalities is given as illustration of the richness of approaches that benefit from AI.

2.2.3.1.   *EEG*

A larger bulk of work dealing with detection of FoG has focused on direct measures of gait, derived from inertial sensors or cameras, but gait measures are mainly used to detect FoG as it happens, while EEG has demonstrated a decent ability to predict an incoming episode a few seconds before its onset. Ardi Handojoseno et al [134] use connectivity measures derived from surface EEG and, by means of a multilayer perceptron, manage to predict FOG with 78% accuracy. They further improve on these results by preprocessing data via directed transfer function and independent component analysis [135]. These results are still achieved offline but show potential for real-time applications that would allow adaptive cueing in devices intended for FoG management and prevention. Scarcity of studies on EEG-based FoG prediction likely stems from the impracticality of setting up EEG equipment in everyday life, but this might change with improvements in the field of mobile EEG devices.

Though predominantly a movement disorder, PD is known to present with non-motor symptoms too, one of which is mild cognitive impairment. Identification of mild cognitive decline is clinically relevant, because it may progress to dementia, but this task is demanding, as it displays wide heterogeneity. Bertouni et al [136] used machine learning algorithms (SVM and kNN) on high density resting EEG data to identify the severity of cognitive impairment in PD patients, splitting them into five groups, with overall accuracy of 84% and 88% for the respective algorithms.

There has been an attempt to diagnose PD from EEG recordings, using state-of-the-art deep learning transformer model borrowed from text processing – BERT [137]. Their protocol included performing finger tapping with both hands for five intervals of 30s. The EEG recordings taken during tapping on 80 PD patients and 24 healthy controls combined with the BERT model showed overall accuracy of 86%.

2.2.3.2.   *Facial analysis*

Quantification of emotional expressiveness could thus help in progression monitoring, but also facilitate research on the aspects of this complex disease that we do not yet fully understand.

Several groups have worked on developing a methodology for automatically analyzing affect and quantifying facial expressivity in PD from video recordings [27], [138]–[141]. They mainly rely on the concept of action units (AU), which refer to the movement of one or more facial muscles, whose different combinations can be used to describe a particular expression. Geometric features are derived from the recordings, and a classification model (or a set of binary classifiers) is then used to recognize action units. A measure of distance from the neutral face can be calculated for each expression [139], and expressivity can be estimated based on measures of intensity, duration and frequency of several types of facial expressive behavior.

3. AIM AND WORKING HYPOTHESES

The main goal of this thesis is to examine the applicability of kinematic analysis of repetitive finger tapping using a system of wearable inertial sensors in discriminating between groups of patients suffering from Parkinson's disease and atypical parkinsonisms, and a control group of patients without neurological conditions, and subsequently to show the usefulness of algorithms of artificial intelligence applied on kinematic signals for individual patient diagnostics. Given that clinical diagnosis of the tested groups of movement disorders is not an easy task, utilizing measurement data and machine learning algorithms could be called upon to aid in this endeavor.

Hypothesis 1. Through analysis of kinematic data collected during the test of repetitive finger tapping, statistically significant differences can be observed between the control group and patient groups suffering from PD, MSA and PSP, as well as differences among the specific disorders.

Hypothesis 2. With the help of artificial intelligence, patients with PD can be discerned on the individual level from persons without neurological disorders

Hypothesis 3. With the help of artificial intelligence, patients with PD and atypical parkinsonisms can be discerned on the individual level

Hypothesis 4. It is possible to programatically choose a subset of relevant features extracted from kinematic signals which increase the performance of classification among the observed disorders

We will also test the applicability of artificial intelligence in analysis of gait, as recorded by a senzorized walkway, testing Hypothesis 2 from a different angle. Gait data will also be used to touch on Hypothesis 4, automatically extracting gait parameters for detection of PD.

3.1 Published scientific papers related to the doctoral thesis

The author of this doctoral dissertation has co-authored the following scientific papers in the area of the doctoral thesis, i.e., computer-aided diagnostics of Parkinson's disease and towards testing the above listed hypotheses:

1. **Belić M**, Bobić V, Badža M, Šolaja N, Đurić-Jovičić M, Kostić VS. Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—A review. Clinical neurology and neurosurgery. 2019 Sep 1;184:105442. **(M23)**
2. **Belić M**, Djurić-Jovičić M, Ječmenica Lukić M, Petrović I, Radovanović S, Popović M, Kostić V, Implementation of discrete wavelet transformation in repetitive finger tapping analysis for patients with Parkinson's disease, IcETRAN, Zlatibor 2016. **(CONFERENCE)**
3. Djurić-Jovičić M, Petrović I, Ječmenica-Lukić M, Radovanović S, Dragašević-Mišković N, **Belić M**, Miler-Jerković V, Popović MB, Kostić VS. Finger tapping analysis in patients with Parkinson's disease and atypical parkinsonism. Journal of Clinical Neuroscience. 2016 Aug 1;30:49-55. **(M23)**
4. Đurić-Jovičić M, Jovičić NS, Radovanović SM, Ječmenica-Lukić M, **Belić M**, Popović M, Kostić VS. Finger and foot tapping sensor system for objective motor assessment. Vojnosanitetski pregled. 2018;75(1):68-77.
5. Miler-Jerković V, Djurić-Jovičić M, **Perović-Belić M**, Ječmenica-Lukić M, Petrović IN, Radovanović SM, Kostić VS, Popović MB. Multiple regression analysis of repetitive finger tapping parameters. In 2014 22nd Telecommunications Forum Telfor (TELFOR) 2014 Nov 25 (pp. 537- 540). IEEE. **(CONFERENCE)**
6. Djurić-Jovičić M, **Belić M,** Stanković I, Radovanović S, Kostić VS. Selection of gait parameters for differential diagnostics of patients with *de novo* Parkinson's disease. Neurological research. 2017 Oct 3;39(10):853-61. **(M22)**
7. Accepted for publication in Heliyon journal as of March 2023: **Belić M**, Radivojević Z, Bobić V, Kostić V, Djurić-Jovičić M, Quick computer aided differential diagnostics based on repetitive finger tapping in Parkinson's disease and atypical parkinsonisms **(M22)**

4. ANALYSIS OF REPETITIVE FINGER TAPPING

This section will present the work aimed at utilization of kinematic signals recorded from the fingers by a cheap, lightweight, and simple to use custom-made system. The goal is to analyze the recordings obtained by participants suffering from PD and atypical parkinsonisms, and rely on artificial intelligence for help in search for parameters that would reflect differences between these disorders. The system should contribute to clinical evaluation by providing a quick diagnostic suggestion in the face of rather similar phenotypes of several degenerative neurological disorders. The chosen solution expands on a repetitive finger tapping test routinely used in assessment of motor decline, as part of the UPDRS-III test battery. The test traditionally relies on the examiner's experience without quantitative measures, whereas the presented system utilizes consumer grade inertial sensors that capture the motion of the fingers during tapping and can provide a form of objective quantification.

The system components are described in Instrumentation. Participant groups and their demographics are given in Participants chapter, followed by a description of the protocol used for testing.

The data obtained through this system will be analyzed with the goal of testing the hypotheses posed in Chapter 3. In Chapter 4.4, statistical differences between finger tapping parameters on the group level will be assessed among four tested participant groups, with the aim of testing Hypothesis 1. We describe the processing steps and statistical analyses used to process and compare parameters of individual taps, then present the obtained results and discuss the findings which point to certain kinematic differences among the tested groups.

Chapter 4.5 describes an approach to discern healthy participants from those with PD using the obtained finger tapping signals, as a test for Hypothesis 2. We describe the analysis based on discrete wavelet transform, feature extraction and classification, then present the results and discuss the findings.

Chapter 4.6 tackles a more demanding task of multi-class classification, aiming to pinpoint differences in kinematic signals among the four participant groups, for the test of Hypothesis 3. Two approaches were used for this purpose – one relying on deep learning, and the other involving traditional machine learning methods. We first reach for deep learning analysis in Chapter 4.6.2, describing the method of classification and data augmentation using variants of convolutional neural networks, and presenting the obtained results. Chapter 4.6.3 dives into the analyses used to transform the obtained kinematic signals, extract and select relevant features and use them as input to a traditional machine learning models for discerning the tested groups. This also tests the Hypothesis 4. Finally, we discuss the obtained results in Chapter 4.6.4.

4.1. Instrumentation

The system used in this study consisted of two IMUs, each containing a triaxial MEMS accelerometer (LIS3DH) and a triaxial MEMS gyroscope (L3G4200, STMicroelectronics, USA) [110], although analysis will be perfomed only on the gyroscope output. One IMU is mounted on the fingernail of the forefinger, and the other on the fingernail of the thumb. Each of the two IMUs are connected via a small flat cable to a sensor control unit attached to the patient's forearm (Figure 13). The signals collected from the SCU are sent wirelessly to a remote personal computer for processing. The PC also has a user-friendly GUI, built in LabView (National Instruments, USA) that enables initiation and stopping of data acquisition, together with real time plotting of the acquired signals. The light weight and small size allow for uninterrupted test performance.

Figure 7 Wearable system used to record kinematic data (Image adapted from [110])

4.2. Participants

The recordings were performed at the Neurology Clinic, Clinical Centre of Serbia, Belgrade. The study was performed in accordance with the ethical standards of the Declaration of Helsinki. All the participants gave informed written consent prior to participation in the study. Fifty-six participants were enrolled from the Movement Disorders Unit at the Clinic of Neurology, Belgrade, including 13 patients with MSA of predominantly parkinsonian type, 14 patients with PD, 16 patients with PSP and 11 healthy controls (HC) with no history of neurological or psychiatric disease. The controls were age- and sex-matched with the overall patient group. For all patients the right-side was predominantly affected by the disorder. Clinical and demographic data are given in Table 2. The groups were compared using one-way ANOVA and Kruskal-Wallis one-way analysis as a non-parametric test. Where indicated, the groups were compared in a pairwise manner using t test or Mann-Whitney U test with Holm-Bonferroni correction for multiple comparisons, or chi square test for categorical variables.

Exclusion criteria included dystonia or any other condition that might interfere with the ability to perform the motor test, a score of under 26 on the Mini Mental Status Examination scale, or smaller than 15 on the Frontal Assessment Battery of tests, a score above 14 on the Hamilton Depression Rating Scale, a history of psychosis or a more serious condition. The patients were tested in the „off" phase and examined by specialists of neurology with abundant experience in treatment of involuntary movements. UPDRS III (Unified Parkinson's Disease Rating Scale) was used to assess the level of motor impairment, as well as disease stage according to the Hoehn & Yahr system.

Table 2 Demographic and clinical features of patients with MSA (n = 13), PD (n=14), PSP (n = 16) and HC (n = 11)

	Age [years]	Gender [F/M]	Disease duration [years]	Hoehn & Yahr stage	UPDRS total	UPDRS III
MSA	58.4 ± 4.8	9/4	3.47 ± 1.5	3.2 ± 0.7	77.2 ± 12.7	45.4 ± 8.6
PD	62.1 ± 9.4	4/10	4.9 ± 4.5	2.2 ± 0.8	48.1 ± 18.7	27.0 ± 9.8
PSP	67.1 ± 8.9	5/11	5.23 ± 2.3	3.8 ± 0.8	79.9 ± 17.2	45.7 ± 10.4
HC	55 ± 8.4	8/3	-	-	-	-
HC-MSA	-	-	-	-	-	-
HC-PD	-	-	-	-	-	-
HC-PSP	p=0.02	-	-	-	-	-
MSA-PD	-	p=0.05	-	p<0.01	-	-
MSA-PSP	-	p=0.02	-	-	-	-
PD-PSP	-	-	-	p<0.001	-	-
Data are presented as mean± standard deviation. P values are given only where significant group differences were found. UPDRS – Unified Parkinson's Disease Rating Scale, UPDRS III - Unified Parkinson's Disease Rating Scale, Part III: Motor Examination, MSA – Multiple system atrophy, PD - Parkinson's disease, PSP – Progressive supranuclear palsy, HC – Healthy controls.						

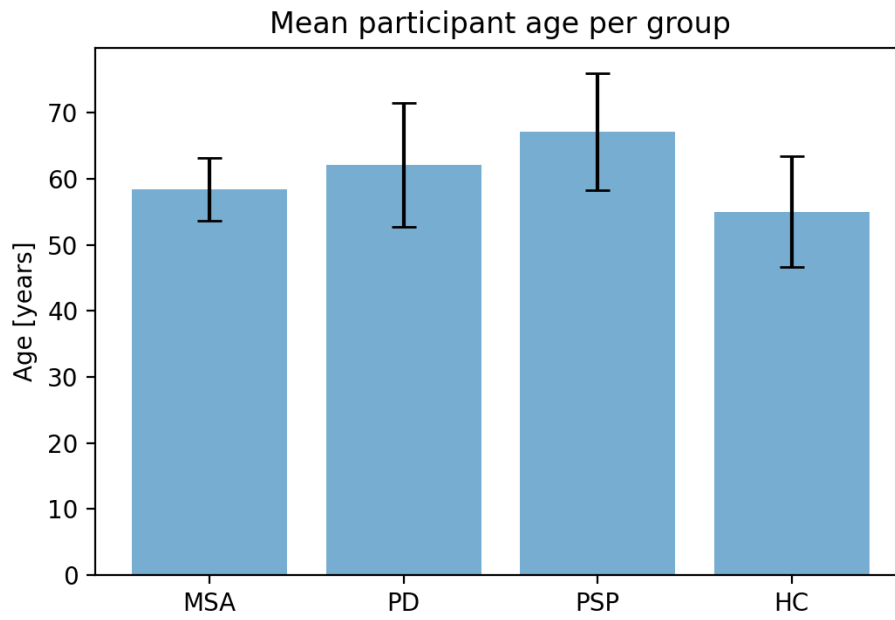Demographic data is shown graphically on figures 7 through 12.

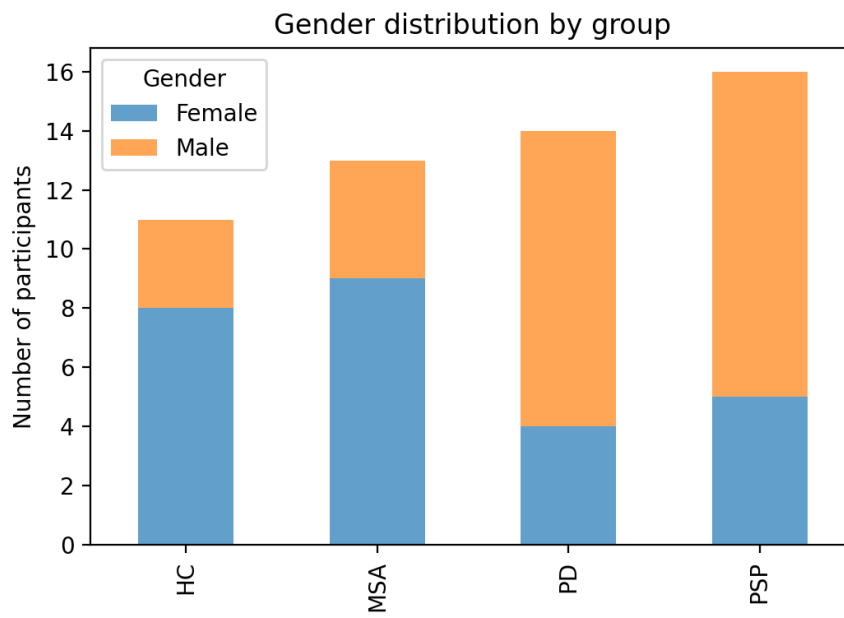*Figure 8 Mean participant age per group*
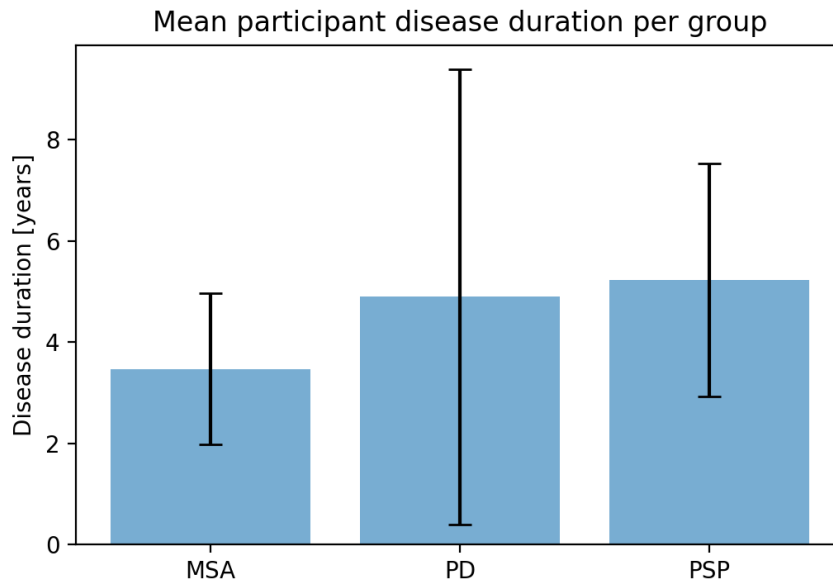

Figure 9 Gender distribution by group

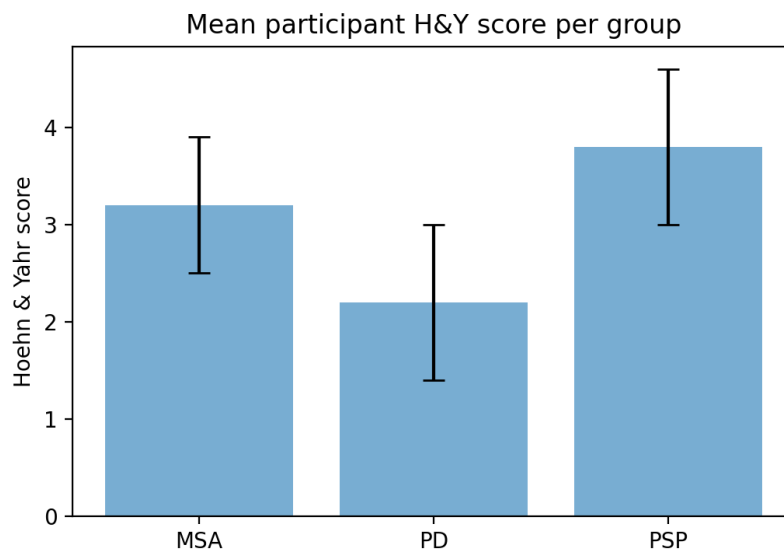*Figure 10 Mean patient disease duration at testing time per group*

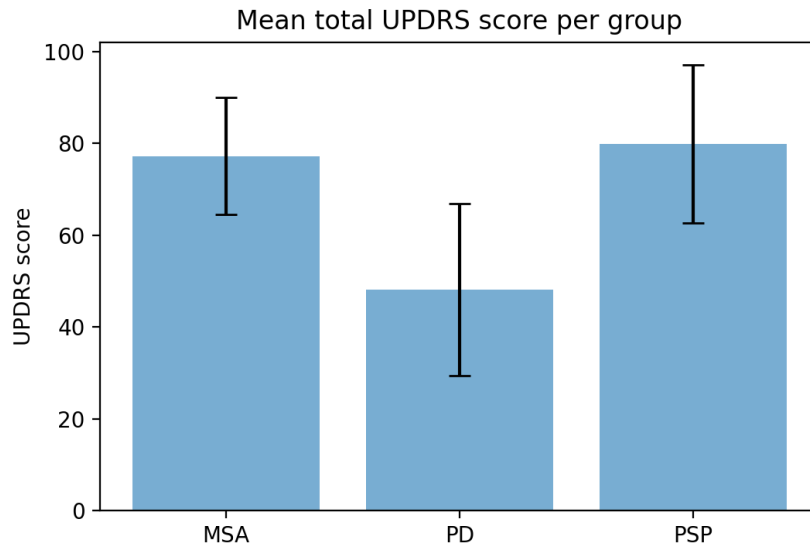*Figure 11 Mean patient Hoehn & Yahr score per group*

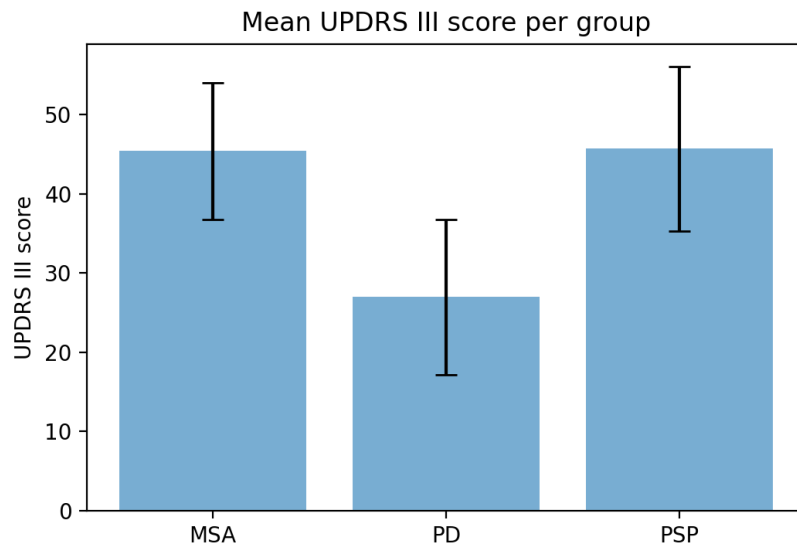*Figure 12 Mean patient UPDRS total score per group*

Figure 13 Mean UPDRS III score per patient group

4.3.   Test protocol

The participants were instructed to sit comfortably and to tap the index finger against the thumb of the right hand repeatedly, as rapidly and as widely as they can for 15 seconds, and a 1 min pause was given between trials. Six trials were taken per patient, although some were later discarded based on data quality and any irregularities that were noticed in the videos that were recorded in parallel. The participants were verbally instructed when to start and stop tapping. Several seconds of variation in duration were allowed, based on experimenter's estimation whether the subject was too fatigued to continue. In the majority of cases, the instructions were understood and followed immediately, since the patients had already been given the traditional finger tapping test for motor assessment prior to the recording session. The patients were recorded on one of their scheduled visits to the clinic, and the total data collection period lasted for a year and a half.

4.4. STATISTICAL ANALYSIS OF BETWEEN GROUP DIFFERENCES IN TAPPING PERFORMANCE

This chapter aims to assess **Hypothesis 1** and find out whether statistical analysis can find significant differences among the finger tapping recordings of the tested participant groups.

4.4.1. Analysis

The gyroscope signals were integrated to obtain the tapping angle. Since this introduces drift error, compensatory correction was performed by finding the beginning of each tap, and then subtracting a 5th order polynomial curve that was fitted through the tap-start points. The stating points were determined through bandpass filtering of the index finger signal using a 4th order Butterworth filter between 0.4 and 5 Hz,after which the signals were squared and peak finding perfomed. Angle amplitude, cycle duration, and speed were measured for each tap, from one separation of index finger and thumb to the next. Tapping amplitude was taken to be the angle between the index finger and the thumb with reference to the long axes. The mean speed was calculated as the mean rate of aperture change, where opening and closing were both taken into calculation. Closing and opening speeds were taken to be the peak velocity of closing and opening within a tap, respectively. Coefficients of variation of speed, amplitude and tap duration over all taps within a recording were also calculated. Linear regression was used to get the best fit for the values of amplitude and speed over time, and the slope of the best fit line was used as a measure of progressive slowing of the tapping motion.

The calculated parameters were compared between groups using ANOVA, or Welch ANOVA for non equal group variances, and Kruskal-Wallis as a non-parametric equivalent. Where significant differences were found, multiple comparisons were done between pairs of groups (Tukey, Games-Howell and Holm test within ANOVA, Welch-ANOVA and Kruskal-Wallis, respectively). Speed and amplitude slopes were compared through univariate analysis of covariance (ANCOVA) with sex, age and disease duration as covariates.

4.4.2. Results

Significant differences were found by ANOVA for all parameters except for the slope of tap duration. The results are summed up in Table 3 below[3].

The largest cadence was found in patients with PSP, although it was only significantly different from MSA, which had the lowest cadence. Statistically significant differences could also be seen between the MSA group and the HC group. Mean tap cycle duration was found not to be meaningfully different from that of the control group, but was notably shorter in comparison with PD and MSA. Analysis of coefficients of variation of tap cycle duration found significant irregularities for PD and MSA groups compared to HC. Duration slopes were not significantly different between HC, PD, PSP and MSA. The higher cadence in PSP, with a relatively stable amplitude (S=-0.12) and speed (S=-0.64) slopes appears to be related to the shorter tap duration.

The mean amplitude was the lowest in the PSP group, while MSA had the highest number among the patient groups, though the differences were only significant between each patient group and HC, but not among the patient subgroups. The HC group had the heighest amplitude overall, together with the smallest coefficient of variation, which was also significantly different from each individual patient

---

[3] An adaptation of the presented results has been published in [143].

group. As for the amplitude slope, PSP (S=-0.12) was similar to HC (-0.21), but the PSP slope was significantly less negative than that of PD (S=-0.56) and MSA (S=-1.48). The significance came to light after adjusting for mean amplitude as a covariate. Overall, the MSA group had the steepest negative slope, as shown in Fig 14. Velocity parameters (mean, opening and closing velocity) were also significantly different between HC and each individual patient group, but did not differ among the patient groups, although the coefficient of variation of opening and closing velocity did significantly differ for PSP and PD, as well as PSP and MSA, with PSP having the lower value, i.e. being more regular. Slopes of the velocity parameters were significantly different between HC and MSA, and HC and PD, but not between HC and PSP. On the other hand, there were significant differences between PSP and MSA, as well as PSP and PD on all velocity slopes, with the PSP slopes being smaller than in the other patient groups. PSP patients also had almost no amplitude decrement over time, reflected in the smallest amplitude slope of all the patient groups.

Table 3 Analysis of kinematic parameters during the finger tapping task in HC and patients with MSA, PD and PSP

Kinematic parameter	HC	MSA	PD	PSP	All groups p value	HC-MSA	HC-PD	HC-PSP	MSA-PD	MSA-PSP	PSP-PD
Cadence [n/15s]	47.8 ± 12.6	27.2±16.9	42.3 ± 18.4	57.6±9.6	p<0.001	p=0.006	-	-	-	p<0.001	-
Duration [ms]	331.7±76.8	808.4±562.6	435.8±211.5	268±53.9	p=0.001	p=0.008	-	-	-	p<0.001	p=0.039
Duration CV [%]	14.5±6.9	24.8±11.4	22.4±6.5	18.8±5.1	p=0.005	p=0.044	p=0.035	-	-	-	-
Duration slope [ms/cycle]	0.04±1.15	9.87±21.03	2.27±5.58	-0.3±1.24	p=0.144	-	-	-	-	-	-
Amplitude [°]	81.8±33.9	40.6±21.2	33.8±11.9	31.4±15.1	p<0.001	p=0.004	p<0.001	p<0.001	-	-	-
Amplitude CV [%]	12.3±5.4	37.2±16.5	32.4±7.8	26.3±7.3	p<0.001	p<0.001	p<0.001	p<0.001	-	-	-
Amplitude slope [°/cycle]	-0.21±0.46	-1.48±1.13	-0.56±0.48	-0.12±0.26	p<0.001	p=0.001	p=0.012	p=0.032	p=0.03	p=0.003	p=0.001
Speed [°/s]	516.6±213.9	143.0±86.5	194.5±91.4	244.8±107	p<0.001	p<0.001	p=0.001	p=0.006	-	-	-
Speed CV [%]	16.1±6.7	36.3±16.0	32.5±8.4	24.4±6.5	p<0.001	p=0.002	p<0.001	p=0.026	-	-	-
Speed slope [°/s/cycle]	-1.88±3.89	-3.99±2.6	-2.89±2.21	-0.64±0.93	p=0.022	p=0.018	p=0.014	-	-	p=0.01	p<0.001
Open velocity [°/s]	1148.1±499	369.6±201.5	458.8±188.9	544.3±206	p<0.001	p<0.001	p=0.002	p=0.008	-	-	-
Open velocity CV [%]	13.2±4.5	34.9±14.7	31.6±8.3	20.9±4.6	p<0.001	p<0.001	p<0.001	p=0.005	-	p<0.001	p=0.001
Open velocity slope [°/s/cycle]	-8.19±12.24	-11.5±11.4	-7.07±4.29	-1.86±2.82	p=0.057	p=0.021	p=0.081	-	-	p=0.002	p=0.003
Close velocity [°/s]	-1602.7 ±503.1	-72.9±256.7	-32.7±287.3	-784.8±346	p<0.001	p<0.001	p<0.001	p>0.001	-	-	-
Close velocity CV [%]	-13.8±5.5	-40.3±15.6	-34.0±9.8	-21.6±3.8	p<0.001	p<0.001	p<0.001	p=0.006	-	p<0.001	p<0.001
Close velocity slope [°/s/cycle]	8.36±9.34	14.84±11.0	10.63±9.65	2.74±3.46	p=0.012	p=0.021	p=0.018	-	-	p=0.003	p=0.013
Data are presented as mean ± standard deviation. Statistical significance is expressed as p values for the comparisons of parameter values. CV = coefficients of variation, HC = healthy controls, MSA = multiple system atrophy of parkinsonian type, PD = Parkinson's disease, PSP = progressive supranuclear palsy.											

Compared to the mean amplitude of the HC group, 86.6% of finger tapping trials of the PSP group had an amplitude that was smaller by 50% (hypokinesia). This percentage was 85% in PD group, and 50% in MSA group. The majority of PSP patients (66%) presented with hypokinesia without decrement, i.e. hypokinesia in combination with the absolute slope being less than 0.1. No such patients were identified in the MSA group, while 27.3% of the PD group corresponded to these criteria.
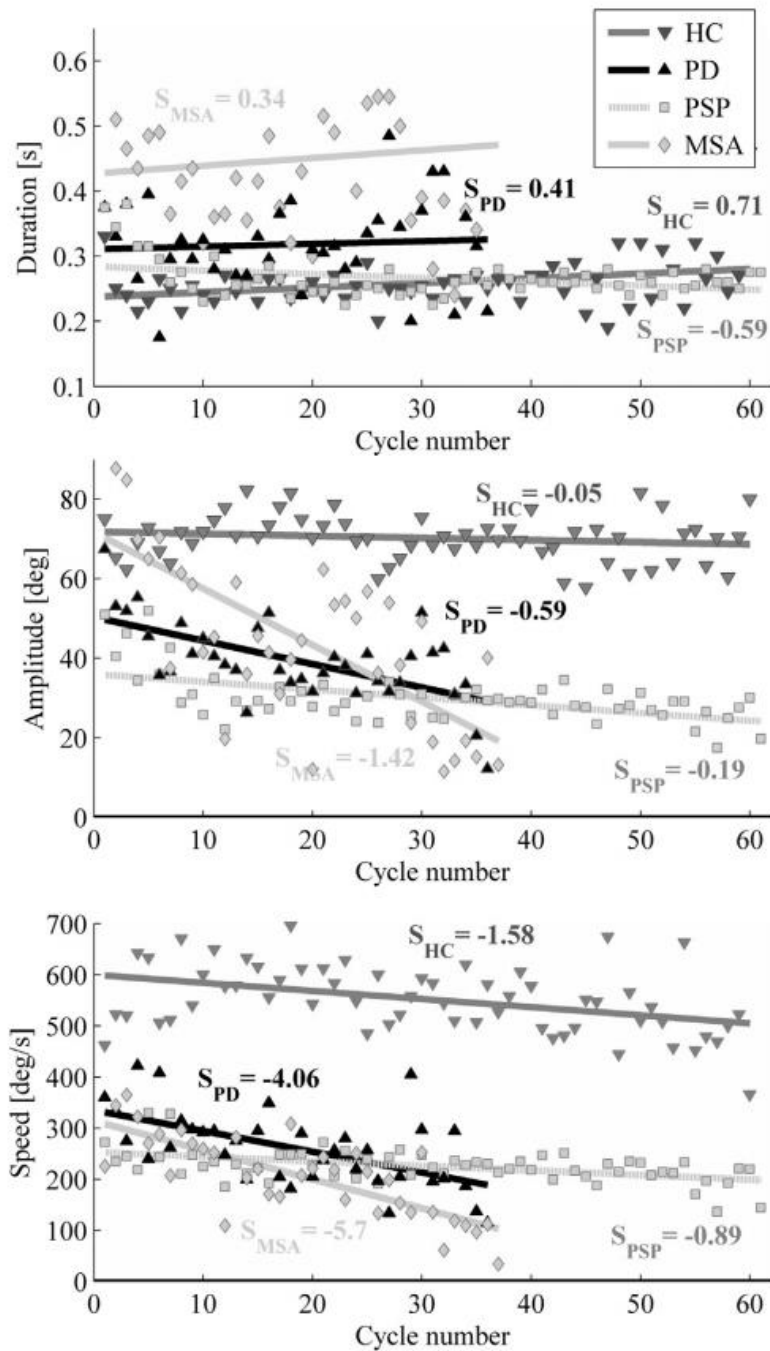
Figure 14 Progression of kinematic parameters over time during a period of 15s, shown for one representative patient per group. Slope was denoted as S, and the linear regression line drawn through the data points (published in [143])

4.4.3. Discussion

Statistically significant differences in certain parameters were found among the tested groups, confirming thus **Hypothesis 1.**

Statistical analysis of kinematic finger tapping parameters showed as the most notable finding the lack of progressive temporal reduction in tapping amplitude for the majority of PSP patients and HC, which was not the case for PD and MSA groups. This was in line with a previous study that found hypokinesia without decrement in PSP patients (87%) but not in PD (12%) [144]. In our study, hypokinesia without or with minimal decrement (abs(S)<0.1) was found in 66% of PSP patients and 23% of PD patients. The mean amplitude slope in the PSP group (-0.12°/cycle) was even smaller than that of HC (-0.21°/cycle), and notably smaller than PD (-0.56°/cycle), and MSA(-1.48°/cycle). This could mean that this temporal progression may be a characteristic of the disease and can potentially be used to discern types of atypical parkinsonisms. Similar patterns were seen for speed slopes, indicative of fatigue during the test. PSP and HC had similar velocity slopes, while the velocity decrement was more prominent in PD (-2.89°/s/cycle) and MSA (-3.99°/s/cycle), compared to PSP (-0.64°/s/cycle).

The mean tapping amplitude was significantly different between HC and each patient group, although not between the patient groups. Contrary to Ling et al [144], we did not find a significant difference in tapping amplitude between PSP and PD groups. The study protocol may have played a role in this, as the mentioned study used recordings acquired from both hands, whereas here we were only dealing with the dominantly affected right hand.

Morphological abnormalities behind the progressive decrement in amplitude and speed of repetitive actions in PSP have not yet been identified, but several options have been proposed, including differences in basal ganglia [145], [146], premotor cortex, supplementary motor area, sensorimotor cortex [114], and the cerebellum [23]. Lee et al. [148] found a connection of the anterior cingulate cortex [147] and the cerebellar inferior semilunar lobule with the severity of this sequence effect in *de novo* PD patients. The cingulate cortex appears to be moderately affected in PD, but not in PSP-R [34].

4.5. CLASSIFICATION: PARKINSON'S DISEASE VS HEALTHY CONTROLS

Hypothesis 2 states that with the help of artificial intelligence, patients with PD can be discerned on the individual level from persons without neurological disorders (healthy controls). To test this hypothesis, we first transform the gyroscope signals of finger tapping using discrete wavelet transform, and then use a neural network to classify PD patients vs healthy controls.

4.5.1. Analysis

In this chapter, discrete wavelet transformation is applied to the gyroscope recordings of finger taping, and used to extract features, which are then passed to the support vector machine algorithm for classification into PD and HC groups. The overall algorithm used is illustrated in Fig. 15 and will be explained in the text below.
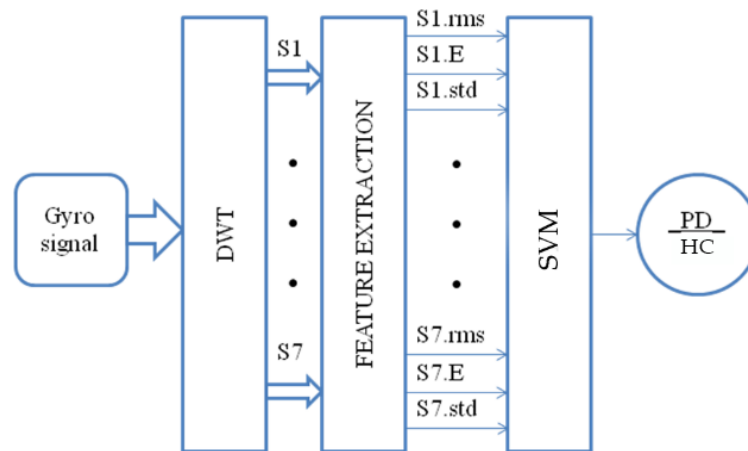

Figure 15 Functional model of the presented algorithm, resulting in classification between patients with Parkinson's disease (PD) and healthy controls (HC) (Published in [162]*)*

Discrete Wavelet Transformation

Wavelet transformation (WT) is a type of time-frequency analysis useful for non-stationary signals, i.e. signals whose frequency content varies over time, which is often the case with biological signals. WT has an advantage over short-time Fourier transform in that it can preserve information with high resolution in time for higher frequencies and frequency resolution in lower frequencies. Apart from its well-established uses in image compression and noise reduction, WT has found application in signal processing for analyzing heart rate variability [149], and discerning healthy individuals from those with cardiac pathology [150]. Several research studies [151]–[153] suggest that the application of discrete WT is suitable for extracting the features that can be used effectively to compare and classify data obtained by inertial sensors. Triaxial accelerometer data recorded during walking [151] were used in combination with a multilayer perceptron neural network (NN) and features extracted with discrete WT (DWT) to recognize 5 different walking patterns, including walking in a straight corridor, up and down a flight of stairs, or up and down a slope. DWT extracted features from leg-mounted gyroscope signals were also used as input to a multi-layer feed forward artificial NN to achieve near perfect classification of leg motion types, such as knee bending, squatting, moving the leg forward or backward and more [154]. Abnormal gait patterns were detected using DWT and a convolutional neural network applied to

data gathered from two inertial sensors in a group of children with cerebral palsy and a group of healthy children. This work showed superior performance of the DWT approach compared to classification without DWT preprocessing [155]. To address threats of falls and postural instability, DWT was used for postural sway classification, achieving 100% sensitivity and 96% specificity [156]. Another variant of wavelet decomposition, namely continuous wavelet decomposition has proven to be an effective method of extracting information from gyroscope measurements for classification of repetitive finger tapping [80], however this type of wavelet transform is more computationally expensive than the proposed DWT.

Wavelet transformation transforms a signal by convolving it with a set of mathematical, basis functions – wavelets, which has the effect of decomposing the input into different components with different frequency contents. Continuous wavelet transform is described by equation (1) [157]:

$$H(x) = \frac{1}{|\sqrt{\zeta}|} \int x(t) \cdot \psi * (\frac{t-\tau}{\zeta}) dt \tag{1}$$

Where H(x) is the wavelet transformation of the input x(t) as a function of time, ζ denotes the scale, τ stands for time, and ψ is the basis function, the mother wavelet, while operation * is the complex conjugate. The scale parameter is the inverse of frequency. The time parameter moves along the signal and gives temporal information. DWT uses sub-band coding to compute the transformation, by effectively filtering the signal through high-pass and low-pass filters, recursively down-sampling the signal by a factor of two, resulting in approximation and detail coefficients, represented by Mallat's decomposition tree (Figure 16) [158] It is then possible to reconstruct the original signal from the wavelet coefficients by up-sampling by two, passing through high and low pass synthesis filters and summing them. Noise reduction can be done by reconstructing only the desired decomposition levels and discarding the others.
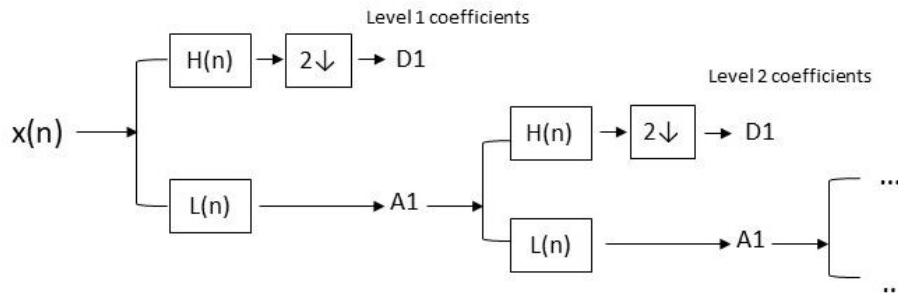

Figure 16 Mallat's decomposition tree showing how a signal gets passed through a high pass filter H, and low pass filter L, and thus decomposed into a coarser resolution approximation A and signal detail D. This is repeated in a cascade up to a desired level of d (Adapted from [158])

For this analysis, signals recorded from the sensors were transformed using discrete wavelet transformation and Daubechies 4 mother wavelet (Figure 17 and Figure 18). The analysis was performed up to the 7th level of decomposition, yielding coefficients that correspond to frequencies from 100Hz to approximately 1Hz. Processing was done using custom-made software written in Matlab (The Mathworks Ltd, USA).
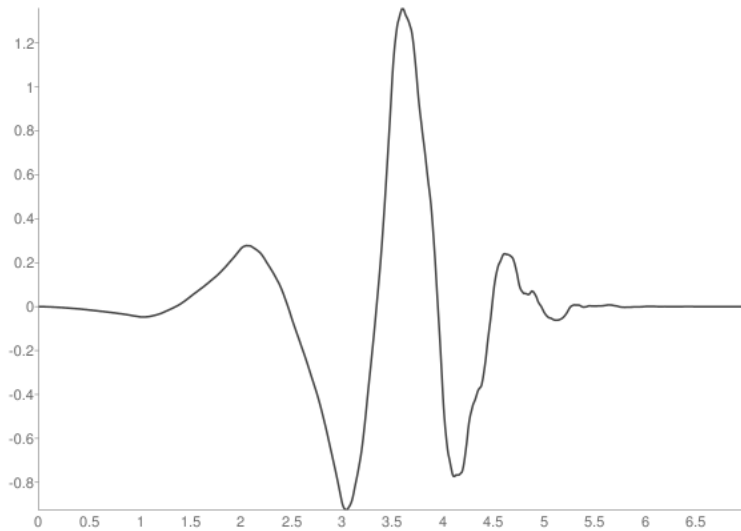
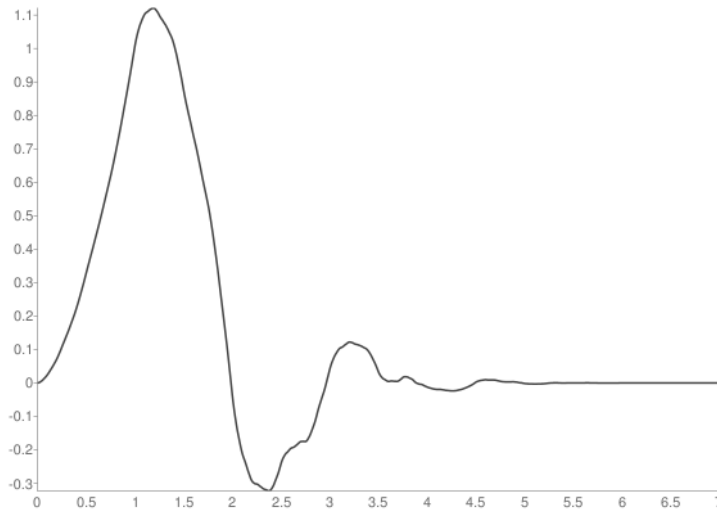*Figure 17 Daubechies 4 wavelet function (Adapted from* [159]*)*

Figure 18 Daubechies 4 wavelet scaling function (Adapted from [159]*)*

The scales and their corresponding frequency bands in the signal are presented in Table 4.

Table 4 Dwt scales and their corresponding frequency bands

Scale	Frequencies [Hz]	
	From	To
1	50	100
2	25	50
3	12.5	25
4	6.25	12.5
5	3.125	6.25
6	1.56	3.125
7	0.78	1.56

Feature extraction

Once the signal was decomposed using DWT, several features were calculated from the obtained coefficients on each scale. This included wavelet energy, root mean square and standard deviation of wavelet coefficients, calculated according to the formulas (2), (3) and (4):

$$We(m) = \sum_{n=1}^{N} D_{m,n}(s)^2 \tag{2}$$

$$RMS(m) = \left[\frac{1}{N}\sum_{n=1}^{N} D_{m,n}(s)^2\right]^{\frac{1}{2}} \tag{3}$$

$$STD(m) = \left[\frac{1}{N}\sum_{n=1}^{N}\left[D_{m,n}(s) - \langle D_{m,n}(s)\rangle\right]^2\right]^{\frac{1}{2}} \tag{4}$$

where N is the number of wavelet coefficients on scale m, and $D_{m,n}$ is the value of DWT details, related to the scale and translation *n*, for each subject *s*. In addition to the above features extracted on the whole signal, the recordings were split into five equal temporal windows (each containing approximately 3s of the finger tapping sequence) and the features described above were calculated from each window. Features were not taken from the first two levels of decomposition, since their corresponding frequency bands were not expected to contain useful information but rather to contain noise. Subtracting the wavelet energy, RMS and STD calculated on the last signal window from the corresponding values obtained for the first window in the signal yielded additional features that were taken into consideration. We will call this group the "delta" features. The inclusion of this group of features was motivated by the observation that characteristics of the disorder can be reflected not only in the statistics derived from the signal as a whole, but also in their temporal progression, as it was found that a progressive reduction in amplitude and speed over repeated sequences is a feature that distinguishes PD from PSP [144]. Difference in means between the control and PD groups was tested using unpaired student's t test for all features, and only those features that showed a significant difference between the two groups at alpha level of 0.05 were kept for further processing. This resulted in 65 features which were then passed to the classifier.

Support vector machine

The extracted features were fed to a support vector machine with radial basis kernel (using the *caret* library in R). To assess the contribution of delta features, classifier was also trained without these features, i.e. it only contained whole signal parameters. The classification goal was to discern PD patients from healthy controls. The accuracy of the classifier was assessed using 4-fold cross validation repeated 10 times.
Sensitivity and specificity were also calculated according to the formulas (5)and (6):

$$Sensitivity = \frac{No\ of\ true\ positives}{No\ of\ actual\ positives} * 100\% \tag{5}$$

$$Specificity = \frac{No\ of\ true\ negatives}{No\ of\ actual\ negatives} * 100\% \tag{6}$$

Sensitivity refers to the ability of the classifier to identify patients with the disease, while specificity represents the ability to correctly identify persons without the disease.
Support vector machine (SVM) is a supervised machine learning method of classification. Its core is the search for the parameters of a hyperplane that would best separate the given data set into two semi-spaces. Le the dataset be given as {$\mathbf{xi}$, yi}, i = 1..m, where where *m* is the number of observations, $\mathbf{xi}$ is

the i-th sample represented as an n-dimensional vector and the labels yi are paired with each **xi** sample, denoting its belonging to either semi space. In our case, **xi** would be a 65-dimensional feature vector and yi would be either -1 or 1, signifying HC and PD respectively. The hyperplane is searched for in such a manner that minimizes the upper bound of the generalization error via maximizing the margin between the decision boundary, i.e. the hyperplane, and the data points closest to the boundary [160]. The data points that are the nearest to the separating hyperplane are termed "the support vectors", hence the name of the algorithm. The SVM algorithm aims to solve the optimization problem of maximizing the margin, which is the projection of the support vectors onto the normal of the hyperplane **w**. This makes the optimization problem equal to the following:

$$\arg min \frac{1}{2} \|w\|^2 \tag{7}$$

This minimization of the hyperplane parameters should be done while under the constraint that [161]:

$$y_n(w \cdot \varphi(x_n) + b) \geq 1 \tag{8}$$

where **w** is a vector of weights or parameters of the hyperplane, φ (x) denotes a fixed feature-space transformation of the known data, and *b* is bias. The term in the brackets corresponds to the separating hyperplane. This equation is the result of the initial aim: if yi is 1 for a given **xi**, then we would like the inclusion of **xi** into the hyperplane expression to be greater than some margin γ, and if yi is -1 then we would like the hyperplane expression for **xi** to be smaller than -γ. It was shown that replacing γ with 1 does not lead to loss of generality. So, by taking the product of the two elements, we get to the equation (8).

Splitting the vector space into two subspaces by a hyperplane, implies that the data are linearly separable, which may not originally be the case. To address this, a mapping function **φ** is chosen to project the data onto a high-dimensional (ideally infinite-dimensional) space $F = R^N$ (N>>m) in which it would be linearly separable. However, performing this would be computationally overwhelming, thus a kernel function *K(x,x')* is employed to transform the data. It can be shown that it now becomes unnecessary to directly calculate the mapped pattern **φ(x)**, but instead only use dot products. The radial basis function (RBF) kernel, or Gaussian kernel, is a common choice for SVM implementations, and is defined as:

$$K(x, x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \tag{9}$$

Where $\|x - x'\|^2$ denotes Euclidian distance between two feature vectors, and σ is a kernel parameter that is tuned for a specific application. Plugging this into the hyperplane expression, the decision function becomes:

$$f(x) = sgn(\sum_1^m \alpha_i y_i \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) + b) \tag{10}$$

4.5.2. Results

A sample of a raw gyro signal recorded from a healthy subject is shown in Fig 19 and its corresponding discrete wavelet decomposition can be seen in Fig 20. The figure shows DWT details for scales 1 through 7 (denoted as S1-S7)[4].
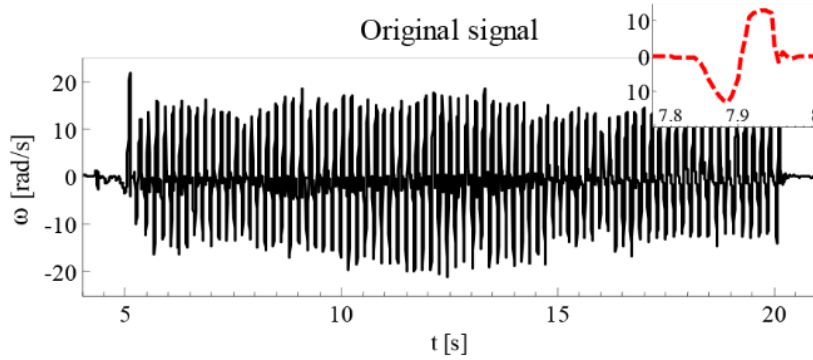

Figure 19 Raw gyroscope signal recorded from the index finger of a healthy participant: full recorded sequence (black solid line) and one isolated tap (red dashed line). (Published in [162]*)*
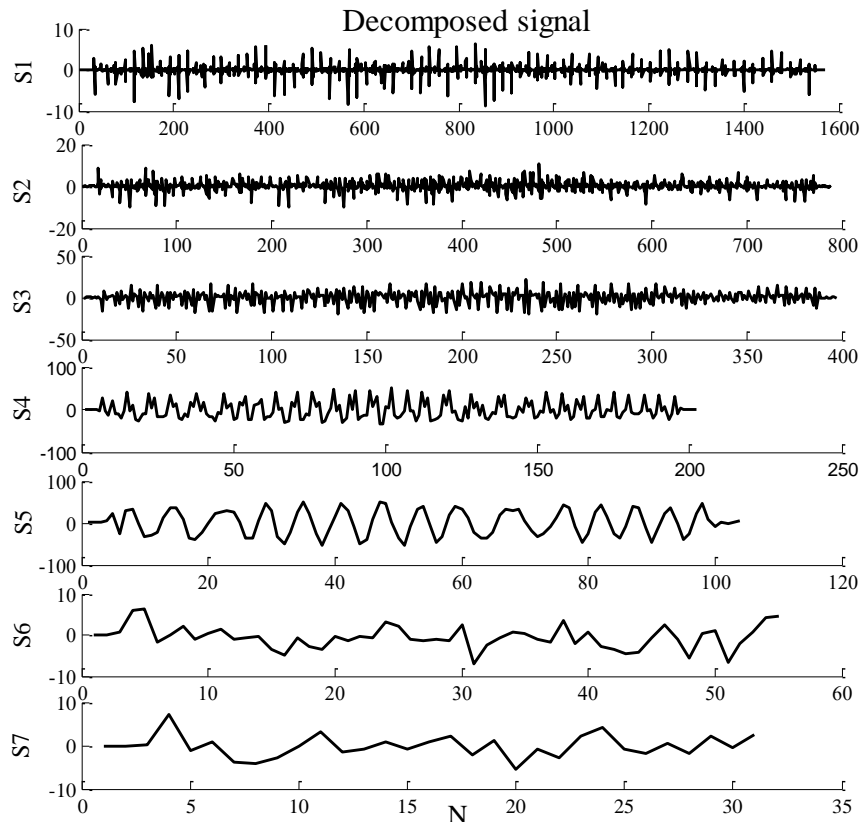

Figure 20 DWT decomposition of the gyroscope signal shown in previous figure (Published in [162]*)*

Wavelet energy, RMS and STD of wavelet detail coefficients extracted for each scale on the whole signal

---

[4] Parts of these results have been adapted and presented at IcEtran conference 2016 [162].

are shown in Figure 21. Different colored lines on each graph shows the calculated features from different subjects in the given group (PD or Healthy participants). The subplots alternately depict results obtained for healthy subjects and Parkinson's disease patients for a particular feature.
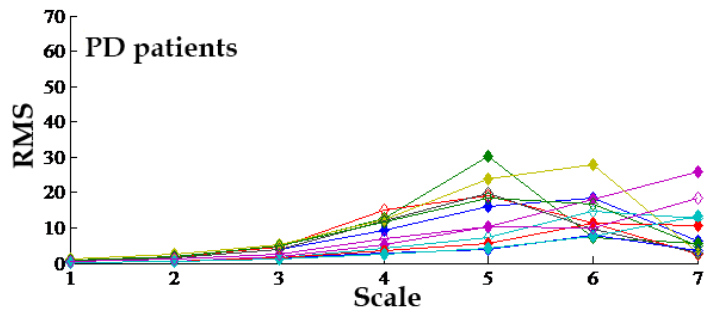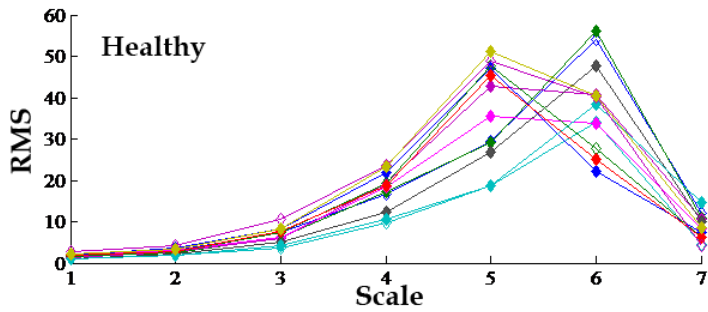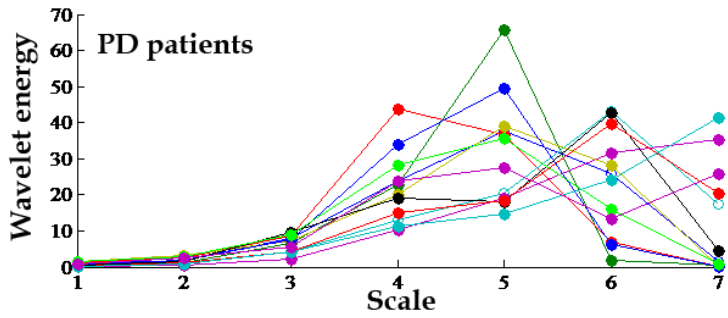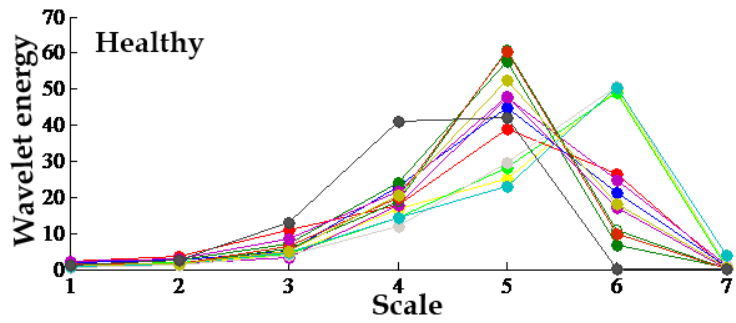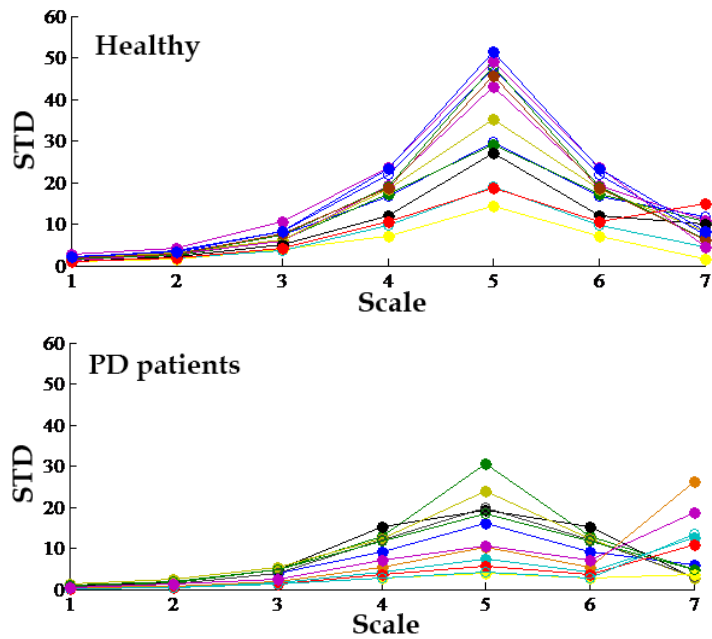
Figure 21 Graphical representation of whole signal features shown alternately for healthy subjects and PD patients (Wavelet energy, RMS, STD) for all subjects in the respective groups (Published in [162]*).*

We see from observing the graphs that wavelet energy is largely focused on the 5th scale (3.125-6.25 Hz) for healthy participants and notably less on the 4th, wavelet energy on the 4th (6.25 – 12.25 Hz) and 6th (1.56 – 3.125Hz) scales is comparable to that on the 5th scale, implicating both lower tapping cadence and high frequency tremor. The 7th (0.78 – 1.56 Hz) scale is nearly completely absent in healthy participants but it quite prominent for the PD group.

RMS coefficient values for healthy subjects are much higher in absolute values (up to 60) than PD patients (not exceeding 40). Peak values in both healthy and PD participants are seen on the 5th and 6th scale, although the peaks for the healthy group are more prominent, and fall sharply on scale 7, while for certain participants in PD group there is a rise in RMS up to the 7th scale.

STD values similarly peak on the 5th scale for both groups, but consistently fall on the 7th for the control group, and rise on the 7th for the PD group, in some cases above all others, suggesting large variability in the presence of low frequency content, suggesting that the tapping is not consistent but is interspersed with slowing or pauses in the sequence.

Figure 22 shows a sample delta feature, the difference between RMS values calculated for the first and the last window of the fifth level of DWT where RMS peaks (deltaRMS5) for all subjects. There is a clear tendency of RMS on this scale to drop over time for PD patient, seen in deltaRMS values being larger than zero, meaning the RMS was higher at start than at the end of the tapping sequence. This is not the case for healthy controls, where this difference stays close below zero or even drops significantly, meaning that in most cases the energy exerted in the end of the tapping sequence is somewhat higher than at the start, showing a sort of a warm-up effect.
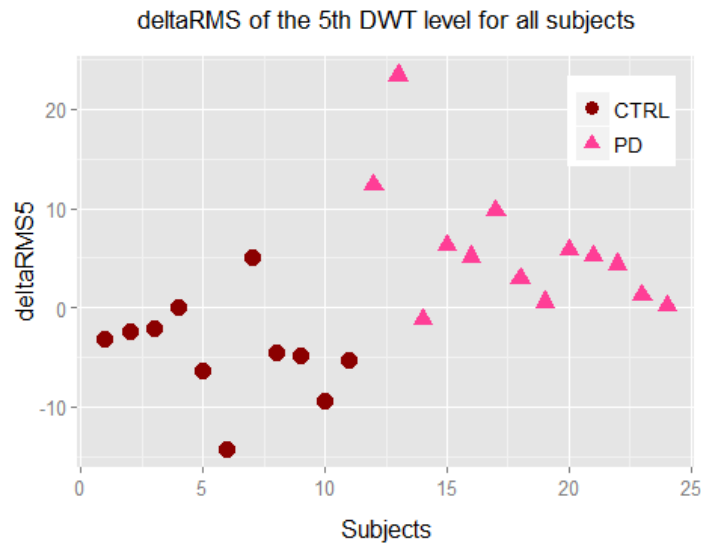
Figure 22 Difference between RMS values on the 5th scale of discrete wavelet decomposition calculated for the first and the last temporal window shown for all subjects (Published in [162])

Feeding the selected features into the SVM RBF algorithm resulted in accurate classification of the data into PD and HC groups with 92% mean accuracy (Table 5) and 11% standard deviation over cross validation trials. Specificity was found to be equal to 100% without variation, meaning that all data points that correspond to healthy participants were classified properly. Sensitivity was 82.5%, which means 82.5% of patients with Parkinson's disease were classified correctly, although this metric had a large variability of 26%.

When the delta features which characterize temporal progression of extracted features were omitted from the classifier, a significant drop in classification accuracy was observed, from 92% down to 86%. Metrics of sensitivity and specificity also dropped, from 82.5% down to 74% for sensitivity and from 100% specificity to 95%, showing positive contribution of the delta features to the algorithm performance.

Table 5 Classification results including accuracy, sensitivity and specificity given as mean±std calculated over cross validation trials

	Accuracy [%]	Sensitivity [%]	Specificity [%]
All features	92±11	82.5±26	100±0
Without delta features	86.1±12	74±29	95±13

44

4.5.3. Discussion

Discrete wavelet decomposition was used to transform gyroscope signals recorded from healthy participants and PD patients. From this transform, 65 features were extracted, some pertaining to the entire signal, and some as a result of subtracting the last fifth of the signal in time from the first one. Some difference in extracted features could be observed between the control and PD groups. For instance, notable is the presence of the 7th level component of wavelet energy for PD patients, and its absence in the control group. Wavelet energy in PD group is generally more dispersed than that of HC group, where it is neatly focused on the 5th scale, corresponding to frequency band 3.125-3.25 Hz, and to a lesser extent the 6th scale (1.56 – 3.125Hz). The values of RMS and STD of wavelet coefficients are largely smaller for PD patients than for the healthy participants. Difference between RMS calculated for the first and the last window of the 5th DWT level has a value that is in almost all cases larger than zero for PD patients and below zero for healthy participants.

Support vector machine with RBF kernel was used in conjunction with the DWT-extracted features and was successful at discriminating healthy subjects from patients with Parkinson's disease, showing overall success rate of 92%, and specificity of 100%, suggesting that all data points corresponding to healthy controls were consistently classified correctly, while mean sensitivity was 82.5%. When comparing the classification model was built on all features excluding the delta features, which denoted temporal differences between the feature values, all obtained metrics had a poorer performance – 86.1% accuracy, 95% specificity and 74% sensitivity. This shows the positive contribution of temporal progression parameters during repetitive finger tapping in identification of Parkinson's disease.

4.6. CLASSIFICATION: MULTICLASS SETTING (HC, MSA, PD, PSP)

Previous chapter showed the usefulness of gyroscope recordings during repetitive finger tapping paired with machine learning in detecting the presence of PD in a pool of healthy participants and patients with PD. This chapter tries to extend these findings to possibly discern not only between healthy controls and PD patients, but also individuals with atypical parkinsonisms, namely multiple system atrophy (MSA) and progressive supranuclear palsy (PSP).

This aims to test **Hypothesis 3**, which states that artificial intelligence can help discern patients with PD and atypical parkinsonisms on the individual level. To this end, we employed two approaches: deep learning and traditional machine learning.

4.6.1. Finger tapping data for multiclass classification

Data was collected from triaxial gyroscopes, at 200 Hz sampling rate in 12-bit resolution. The data set consisted of 268 recordings: 52 recordings from healthy controls, 72 recordings from MSA patients, 68 from PD patients and 76 from patients with PSP. Example signal segments of 4s long raw gyroscope are given as illustration for each patient group in Figure 23.
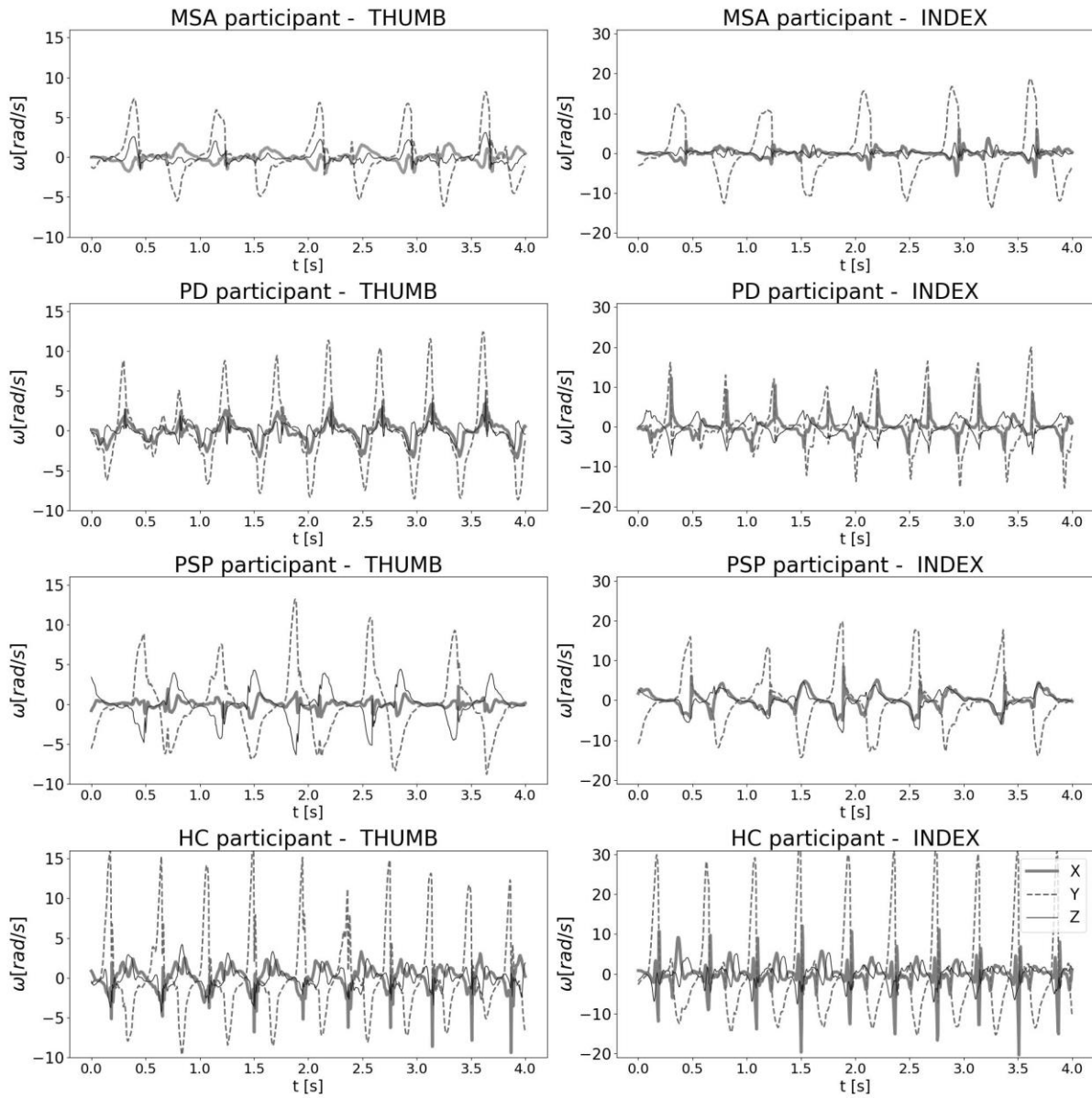
Figure 23 Example gyroscope signals recorded for each participant group. Angular velocities from the thumb gyroscope are presented in the left-hand column, whereas the index gyroscope signals are shown in the right-hand column.

4.6.2. Analysis - Deep learning approaches

4.6.2.1. Classification

Given the immense success of deep learning approaches – recently the most recognizable and celebrated representatives of artificial intelligence - when used on neuroimaging data, we anticipated it would yield superior results in case of tapping signals too. The main concept was to feed the raw gyroscope data into a convolutional neural network, resembling the classic LeNet 5 architecture [163]. Convolutional neural networks have shown the most astonishing results in computer vision problems but have also been used on inertial motion data. For instance, in gesture recognition, which typically uses data from inertial motion sensors, this is the most commonly employed method [164]. The network used here performs 1D convolutions on each of the 6 channels (3 axes from two gyroscopes) through 6 convolutional layers, each followed by batch normalization and a max pooling layer. The network topology is shown in Figure 24.
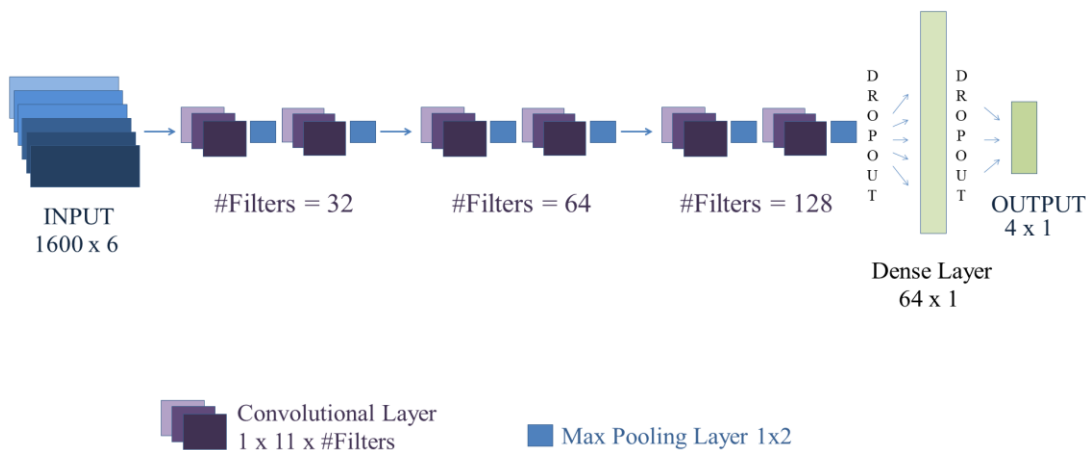

Figure 24 Convolutional neural network containing six layers of one-dimensional convolutions, with the number of filters increasing with depth from 32 to 128. Input data contained 6 channels of 8s long raw gyroscope data sampled at 200Hz.

In recent works, batch normalization has been frequently relied on, since it not only speeds up the complex modern architectures but also acts as a regularizer [165]. It refers to normalizing each input dimension of a layer in a mini batch so that it has a mean of 0 and variance of 1[166]. Max pooling is a pooling operation commonly used after a convolutional layer, in which a kernel of size 2x2 (or larger) is moved in strides of 2 (or other) over the feature map selecting and keeping only the maximal value in that patch. This has the effect of down-sampling the layer output, as well as stabilizing the classification slightly as small translational changes to the image will not significantly affect the outcome. The number of filters used in a single layer increased from 32 in the first two layers, through 64 in the second two layers to 128 in the last two layers. Dropout [167] was added prior to introducing a fully connected layer of 128 units and once again before the output layer (dropout rates of 0.5 and 0.7 respectively). Li et al [168] show that dropout and batch normalization have a detrimental effect if used simultaneously, more

specifically, if dropout layers are found anywhere before batch normalization layers. Therefore, in the presented model no dropout layers are found before the last batch normalization. The weights in all layers of the model had their norm constrained to a value that was tuned as one of the hyperparameters.

The weights were initialized randomly, using He normal initialization [169] as it was expected to be more compatible with here used rectified linear unit activations, compared to another common initialization suggested by X. Glorot [170]. Initialization by weights obtained via a pretrained autoencoder was also tested [171]. The encoder part corresponds to the convolutional part of the classifier network, while the decoder is an inverse of it, where convolutional layers were replaced by deconvolutions, so that the input signal is matched to itself, teaching the model to learn the most important features for signal reconstruction. This method of initialization is sometimes used when a significant amount of data is available, but only a fraction of it is labelled. The autoencoder is then trained on all available data, but classification is performed using the labelled subset only. The use of autoencoders for pretraining has been found to contribute to classification accuracy even when the same data is used for pre-training and classification training, as the autoencoder is expected to summarize the information in the input, as it is forced to recreate it from a small dimensional aggregate. Medical image analysis has seen benefits of this sort of pretraining on the same dataset [172]. High-level representation of the used autoencoder is shown in Figure 25. For validation of succinct representation learning, the flattened layer at the end of the encoder part was used as input to an unsupervised algorithm, namely k-means clustering [173], which assigns each datapoint to a specific cluster based on the shortest Euclidean distance from the cluster centroids. If the most important characteristics of a disease are distilled during pre-training, clustering should be expected to roughly divide the input data into 4 disjunct clusters.
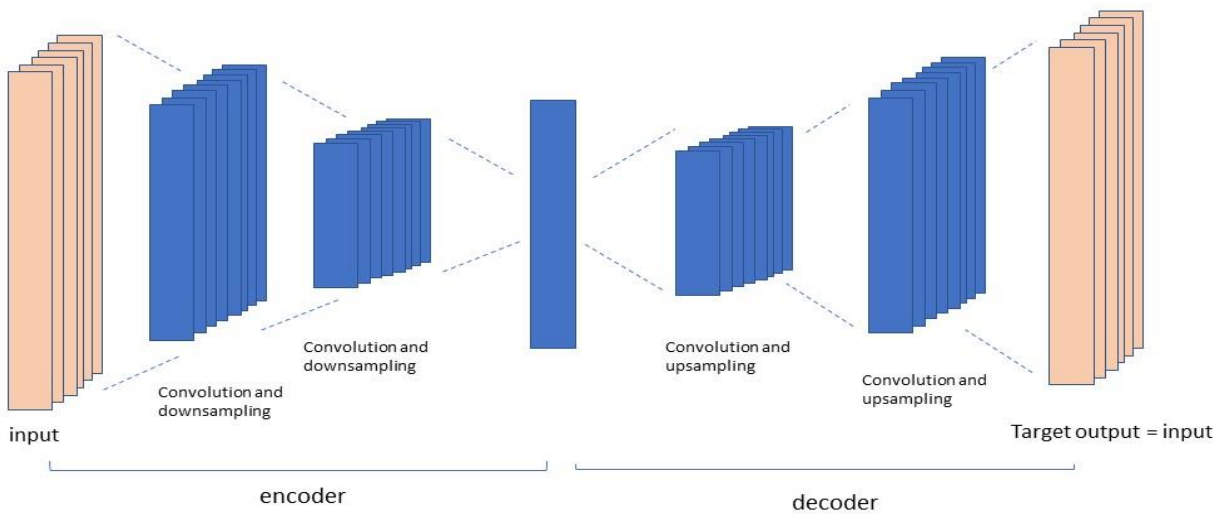

Figure 25 Autoencoder was used for pre-training the classification model. High-level representation

Adaptive moment estimation optimizer (Adam) [174] was used for training, with initial learning rate of 0.001. This nominal learning rate was decayed by a factor of 0.1 once the validation accuracy plateaued for 8 epochs. The model was trained with early stopping, patience set to 20, allowing the learning rate to be decimated twice before abandoning further training. Categorical cross-entropy was taken as the loss function.

Hyperparameters were optimized using the grid search approach.

Apart from feeding raw unprocessed gyroscope data to the network, the signals were also transformed into spherical coordinates using the standard formulae:

$$R \;=\; \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2} \tag{11}$$

$$\theta \;=\; arccos\!\left(\frac{\omega_z}{\sqrt{\omega_x^2+\omega_y^2+\omega_z^2}}\right) \tag{12}$$

$$\varphi \;=\; \arccos\!\left(\frac{\omega_x}{\sqrt{\omega_x^2+\omega_y^2}}\right) \tag{13}$$

Such transformations could possibly make classification more robust to possible changes in hand orientation during tapping.

The model was trained on the training set with the validation set used to assess the generalization efficacy and to guide premature stopping of network training to avoid overfitting. The test set was used for assessment once the best topology had been chosen based on performance achieved on the validation set. This was repeated 5 times, in 5-fold cross validation paradigm, each time taking a different subset of data for training, while the remaining part was held out for testing. Model accuracy was calculated as the number of test instances classified correctly divided by the total number of test instances, expressed in percentages, and averaged over the 5 folds.

4.6.2.2. Data augmentation

Given that deep learning approaches notoriously require large datasets to train, data augmentation was attempted by means of generating synthetic data based on the existing dataset, using a generative adversarial network (GAN). GANs are dual neural networks consisting of the generator and the discriminator (sometimes called critic), where two agents compete against each other. The generator's aim is to use a random noise vector to create synthetic data that would serve as input to the discriminator, alongside genuine data, and fool the discriminator into recognizing generated data as genuine, whereas the discriminator's aim is the opposite – to learn to better distinguish between genuine and synthetic inputs. The two agents are trained concurrently, ideally at the same rate. GANs have been used in image synthesis, semantic image editing, style transfer, image super-resolution and more [175]. Synthetic data generated via GANs was successfuly used to improve segmentation of CT [176] and MRI images [177], by augmenting the available medical data set, which is usually hard to come by and costly to label.

In this study, the input noise dimension was varied exponentially between 128 and 2048. The generator consisted of a number of generator blocks, the number being dependent on input dimension and other hyperparameters used, each comprising a transposed 1D convolution layer, batch normalization and linear rectifier unit (ReLU) activation, with the exception of the last layer which only contained a transposed convolution. Finishing up with a tanh activation combined with scaling the genuine signals in the range [-1,1] was also tested. The number of blocks was varied to find the best fit, alongside kernel sizes and stride. Label smoothing was added to discriminator labels to prevent mode collapse, which involved randomly assigning a number between 0 and 0.3 to fake signals and 0.7 and 1 for genuine signals.

The discriminator was a fully convolutional neural network, consisting of five convolutional blokcs, each containing a 1D convolutional layer, batch normalization and ReLU activation, except for the last layer that only contained a convolution. Kernel sizes were varied to find the best fit. High level topology of the GAN is shown in Figure 26.

Wasserstein loss with gradient penalty was used for training, as this was shown to help prevent trouble with converging and lead to more stable training [178].
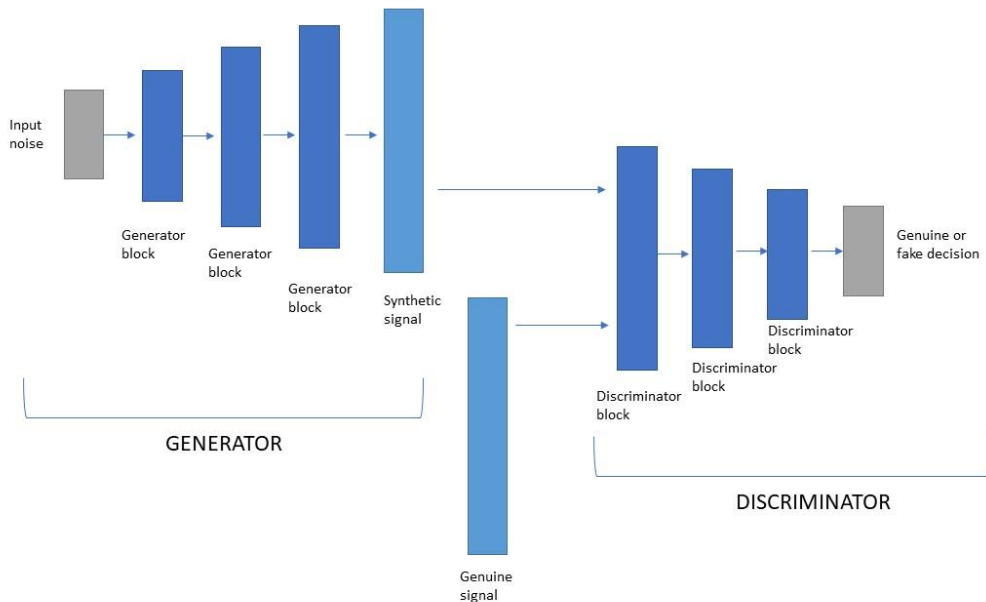

Figure 26 High level diagram of the GAN used for data augmentation

4.6.2.3.   Results

Employed convolutional network model achieved unsatisfactory results, with accuracy on raw signals of 41.5% and 33.1% on spherical transform of the signals. Varying hyperparameters of the network, such as number of layers, number of filters per layer, kernel size, or rate of dropout resulted in no significant improvement of the results. There was a negative correlation (rho = -0.56) of kernel size and accuracy, meaning that larger kernels are less appropriate for this purpose. This may be due to the summation over the kernel destroying potentially useful information in higher frequency details. There was also a negative correlation (rho = -0.43) between utilized training batch size and accuracy, wich implies that larger variance in small batches contributed to model regularization. There was no correlation between the network depth reflected in the number of layers and accuracy, nor between accuracy and the number of neurons present in the dense classification layer of the network nor the number of filters used per layer. Initializing weights with pre-trained autoencoder did not improve the results either.

The trained convolutional GAN exibited failure to converge, irrespective of hyperparameters and tweaks to model topology, and in spite of measures taken to prevent this. The generator and the discriminator loss would fall briefly initially, after which it would stagnate, oscillate, or increase (Figure 27).
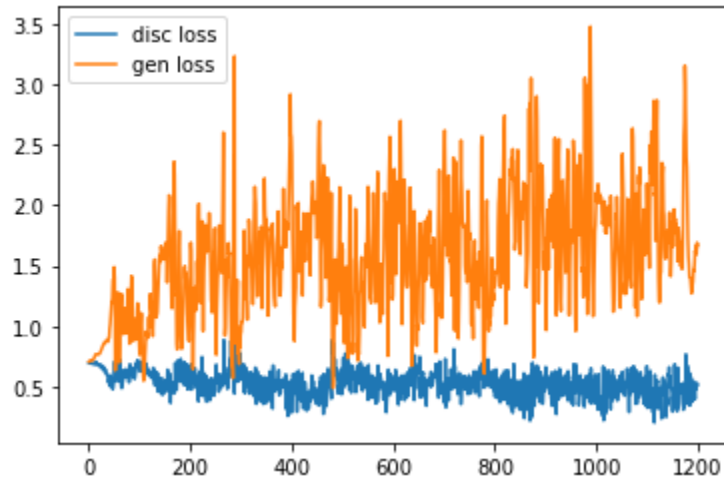
Figure 27 Discriminator (disc) and generator (gen) loss over epochs.

The images generated in the process did not sufficiently resemble the original data, although they did display periodic behaviour. Furthermore, the network would fall into mode collapse, generating only one type of signal at a time, as illustrated in Figure 28 and 29. The synthetic data could therefore not be used to supplement the original dataset.

Figure 28 Example of mode collapse during GAN training, random initialization 1

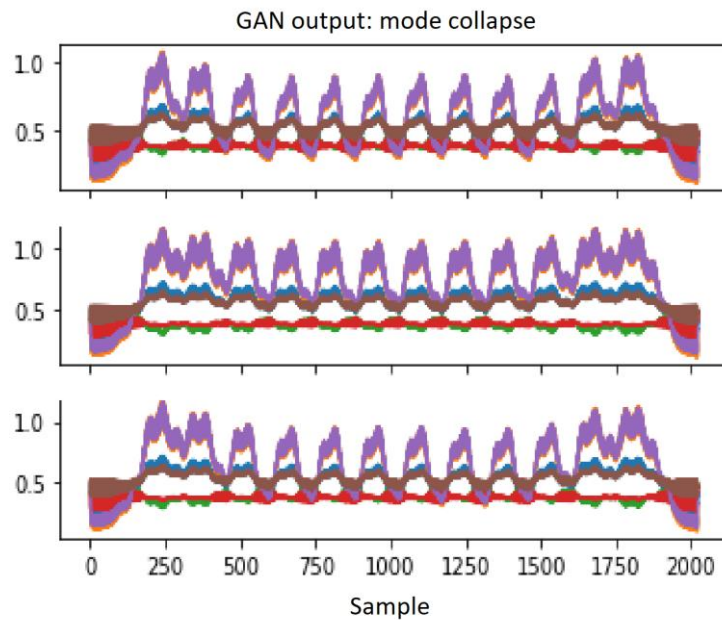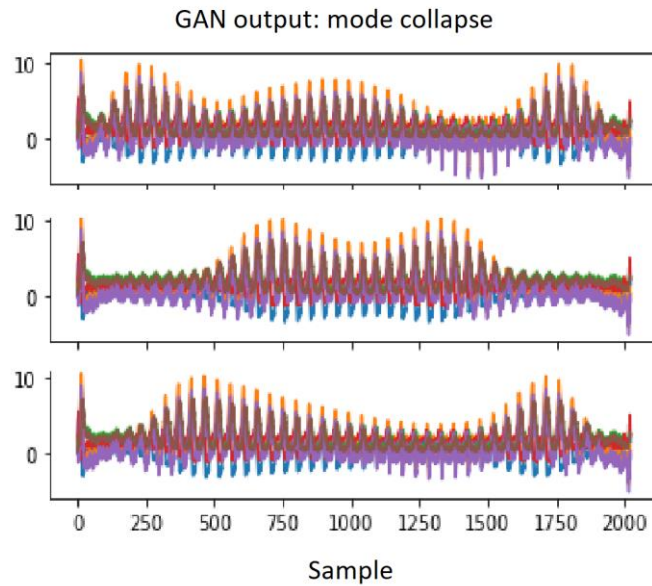*Figure 29 Example of mode collapse during GAN training, random initialization 2*

It is most likely that the available data was insufficient for the deep learning approach, which was reflected in both the inability to obtain a good prediction on the original data, and the failure to generate satisfactory synthetic data to augment the initial dataset. This was also seen in the inability of the autoencoder based clustering to divide the data into meaningful clusters, as the clusters obtained had very low homogenity, suggesting that the autoencoder did not manage to capture the features that were most representative of each disease. Deep neural networks normally require hundreds to thousands of unique samples, and the variance in the existing data combined with the low amount of data did not allow the models to learn the correct mapping functions.

The performed tests were unable to confirm nor deny **Hypothesis 3**. What is clear though, is that the number of tested patients in this study was not enough for deep learning approaches to perform their best.

4.6.3. Analysis - Traditional machine learning approaches

Another branch of artificial intelligence are traditional machine learning algorithms. Unlike deep learning, which usually accepts raw or minimally pre-processed input data, and is relied upon to automatically extract useful features, traditional models accept features that were specifically designed by the researchers. In this chapter, we had a dual aim - to assess **Hypothesis 3**, and utilize traditional machine learning models to discern among groups of patients and healthy controls in a multi-class setup, as well as to test **Hypothesis 4**, and use artificial intelligence to choose the best set of features for discriminating among the participant groups. Although the feature extraction will be done by known formulas given by the researchers in advance, the selection of the best subset of those features will be left to an algorithm[5].

4.6.3.1.   Feature extraction

The right hand signals collected from the participants were processed by handling each axis separately, as well as calculating the norm of all three sensor rotation axes to represent the intensity. The signals were subjected to the following five transformations: 1) no transformation (raw signal), representing angular velocity, 2) integration, yielding the finger excursion angle, 3) differentiation, resulting in angular acceleration, 4) squaring, representing the power of the signal, and 5) Fourier transform, giving the signal frequency content. To compensate for the drift error introduced by integration, the beginning of each tap was determined, and a 5th order polynomial curve was fitted through the tap start points and this curve was subtracted from the integrated signal. The approach to finding tap start times was signal filtering with a 4th order bandpass Butterworth filter, with pass frequencies between 0.4 and 5 Hz. The filter parameters were decided on empirically. The filtered signal was squared to enhance the large changes and suppress the smaller ones, and the last step was negative peak finding. All tap-start positions were then manually checked and corrected where necessary, using a Python script that displayed a visualization of each recording with the automatically determined splits, and allowed user input to remove or add tap start points. (Example tap given in Figure 30).

---

[5] Parts of the presented analysis have been adapted to be published in Heliyon as of March 2023 (Belic M, Radivojevic Z, Bobic V, Kostic V, Djuric-Jovicic M, Quick computer aided differential diagnostics based on repetitive finger tapping in Parkinson's disease and atypical parkinsonisms)

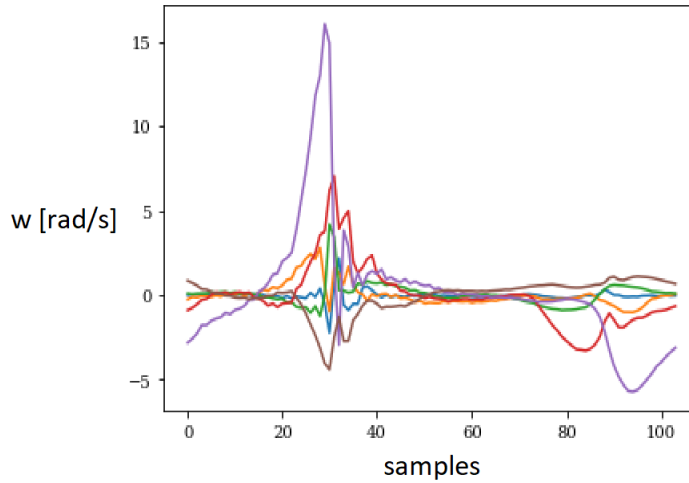Example tap extracted from a recorded gyroscope sequence

Figure 30 Random example of a single tap extracted from a recorded gyroscope sequence

Six features were extracted from each transformed signal, apart from the Fourier transform. These features included the mean, standard deviation, median, root mean square (RMS), and minimal and maximal values. Besides calculating the statistics from the whole signals, the same features were also extracted individual taps, or more specifically from their maximal values. Linear regression line was fitted through these individual taps' maxima, and the slope of the obtained line was taken as an additional parameter. This was to capture the potential amplitude decrements in the tapping sequence over time. Fourier transform was used to extract the maximal frequency and the spectral centroid, which was defined as the frequency point that divides the spectrum into two halves, so that the frequencies below this point account for half of the total spectral power. In total, 216 features were obtained per sensor. The features were then scaled to the range between 0 and 1, as this is expected to help the machine learning algorithms. A summary of extracted features is given in Table 6.

Table 6 Preliminary features extracted from tapping signals

Transformation	Whole signal	Individual taps
No transformation (raw signal)	mean, std, RMS, min, and max of the signal	mean, std, RMS, min, max and regression slope of tap maxima
Integration		
Differentiation		
Squaring		
Fourier transform	max frequency, spectral centroid	

* Std = Standard deviation, RMS = root mean square, min = minimum, max = maximum

55

The number of features was reduced through a semi-greedy feature selection algorithm (Appendix, Algorithm 1). Model classification accuracy was used as a performance measure, and the fitness of a subset of features was rated according to accuracy. The algorithm relied on the broad semi-greedy concept [179]. A full greedy approach would always take the locally optimal solution and add more features in a way the current set of features is the best possible up to that point. In the semi-greedy concept, a broader set of options in allowed, and some randomization is introduced, acknowledging that locally optimal set of features may not be globally optimal too. The algorithm consisted of an expansion phase and a reduction phase, which were repeated multiple times. The expansion phase would generally speaking expand the current feature set but would more precisely modify the current best set by either adding, subtracting, or substituting a feature in a random manner. Whether to use the add, subtract, or modify operation was also decided at random. This was repeated M times, where M was taken randomly to be 2, 3 or 4. Then K feature sets with best performance were saved (K was chosen at random to be 5, 6 or 7) and their union was taken to be the input for the reduction phase. In the reduction phase the aim was to make the feature set smaller, which was done by subsetting and testing all subset combinations of features for their accuracy. The one that gave the best results was then singled out. This process was scheduled to run for 24h on an Intel® Core™ i9 CPU operating at 3.10GHz with 64 GB of RAM.

4.6.3.3.   Classification

In this section, artificial intelligence reflected in traditional machine learning algorithms was used on top of the extracted features to classify the collected recordings into 4 diagnostic groups. A k-nearest neighbors (kNN) classifier [180] was found in preliminary tests to perform the best out of the tested set of algorithms that included random forest, naïve bayes, logistic regression and support vector machine. In kNN, a feature vector is assigned to a specific class, based on the plurality vote of the k vectors in the available training data that are the nearest to the one in question, as measured by the Euclidian distance. Each neighbor point's contribution is weighted by the inverse of its Euclidian distance to the current point. In this work particularly, k is chosen to be 5, that is the class of each recording is determined based on 5 nearest neighbors. Furthermore, since multiple recordings were obtained for each patient, the final decision for each patient was decided as the diagnosis that was predicted most frequently among all corresponding patient's recordings.

4.6.3.4.   Model evaluation

To assess the performance of the model, leave-one-out paradigm was. This involved training the model on all but one patient, making sure that all recordings belonging to that particular patient are held out for the test set. The same trained model was then used to predict the diagnosis for the patient that was left out. This was done for each participant and the results were aggregated to obtain overall performance.

Overall accuracy, precision and recall were calculated for each diagnosis according to the formulas (1), (2) and (3).

$$Accuracy = \frac{Number\ of\ correctly\ classified\ instances}{Number\ of\ all\ instances} * 100\% \qquad (14)$$

$$Precision_i = \frac{Instances\ correctly\ classified\ as\ class\ i}{All\ instances\ classified\ as\ class\ i} * 100\% \qquad (15)$$

$$Recall_i = \frac{Instances\ correctly\ classified\ as\ class\ i}{All\ instances\ of\ actual\ class\ i} * 100\% \qquad (16)$$

For example, if we take the PD group, then precision describes how many of all cases that were suggested to have PD belong to the patients that were clinically diagnosed as PD, while recall describes what percentage of people who were clinically found to have PD will be correctly assigned the PD diagnosis by the algorithm.

All modeling was done using Python 3.7.7 (Python Software Foundation) and scikit-learn package version 0.22.2.

After feature extraction, the semi-greedy feature selection algorithm yielded six features listed in Table 7 as the best combination based on achieved classification results.

Feature 1 is a measure of how much the index finger leans left and right from the path on average. Features 2 and 3 provide information on the sustained finger opening during tapping, with the y-axis rotation being the most dominant one during the finger tapping test. Feature 4 gives insight into the frequency content of the signal, and indirectly the tapping cadence. Feature 5 stands for the maximal rotation of the thumb to left or right, and Feature 6 captures the variability of the thumb's excursion to left and right.

Selected features per group can be found in Table 8, given as the mean value and inter-quartile range.

Table 7 Selected feature set

Feature	Sensor position	Axis of rotation	Signal transformation	Statistic
Feature 1	Index finger	*X*	Angular velocity	RMS
Feature 2	Index finger	*Y*	Angular velocity	average of maxima of individual taps
Feature 3	Index finger	*Y*	Angular acceleration	RMS
Feature 4	Index finger	Vector	Fourier transformation of angular velocity	spectral centroid
Feature 5	Thumb	*X*	Angular velocity	maximum
Feature 6	Thumb	*Z*	Angular acceleration	STD of the maxima of individual taps

* RMS - Root Mean Square, STD - Standard Deviation, vector $=\sqrt{x^2 + y^2 + z^2}$

Table 8 Values of features per group given as mean (interquartile range)

Feature	HC	MSA	PD	PSP
Feature 1 $[\frac{rad^2}{s^2}]$	7.38 (3.89 - 9.87)	0.98 (0.26 - 1.07)	1.96 (1.03 - 2.59)	2.27 (0.59 - 3.50)
Feature 2 $[\frac{rad}{s}]$	21.06 (16.50 - 24.94)	8.74 (3.74 - 13.15)	13.24 (8.31 - 15.32)	11.42 (7.31 - 14.43)
Feature 3 $[\frac{rad}{s^2}]$	2.91 (2.47 - 3.38)	0.77 (0.25 - 1.21)	1.46 (0.68 - 1.93)	1.39 (0.92 - 1.58)
Feature 4 **[Hz]**	12.39 (10.98 - 13.72)	9.11 (7.66 - 10.18)	9.81 (7.68 - 11.49)	11.35 (8.78 - 13.84)
Feature 5 $[\frac{rad}{s}]$	0.60 (0.40 - 0.71)	0.20 (0.08 - 0.23)	0.51 (0.24 - 0.81)	0.26 (0.13 - 0.37)
Feature 6 $[\frac{rad}{s^2}]$	0.89 (0.57 - 1.22)	0.43 (0.16 - 0.64)	0.86 (0.53 - 0.1.21)	0.67 (0.23 – 1.12)
* HC – Healthy Controls, MSA – Multiple System Atrophy, PD – Parkinson's Disease, PSP – Progressive Supranuclear Palsy				

The distribution of features over groups is shown graphically in Figure 31.
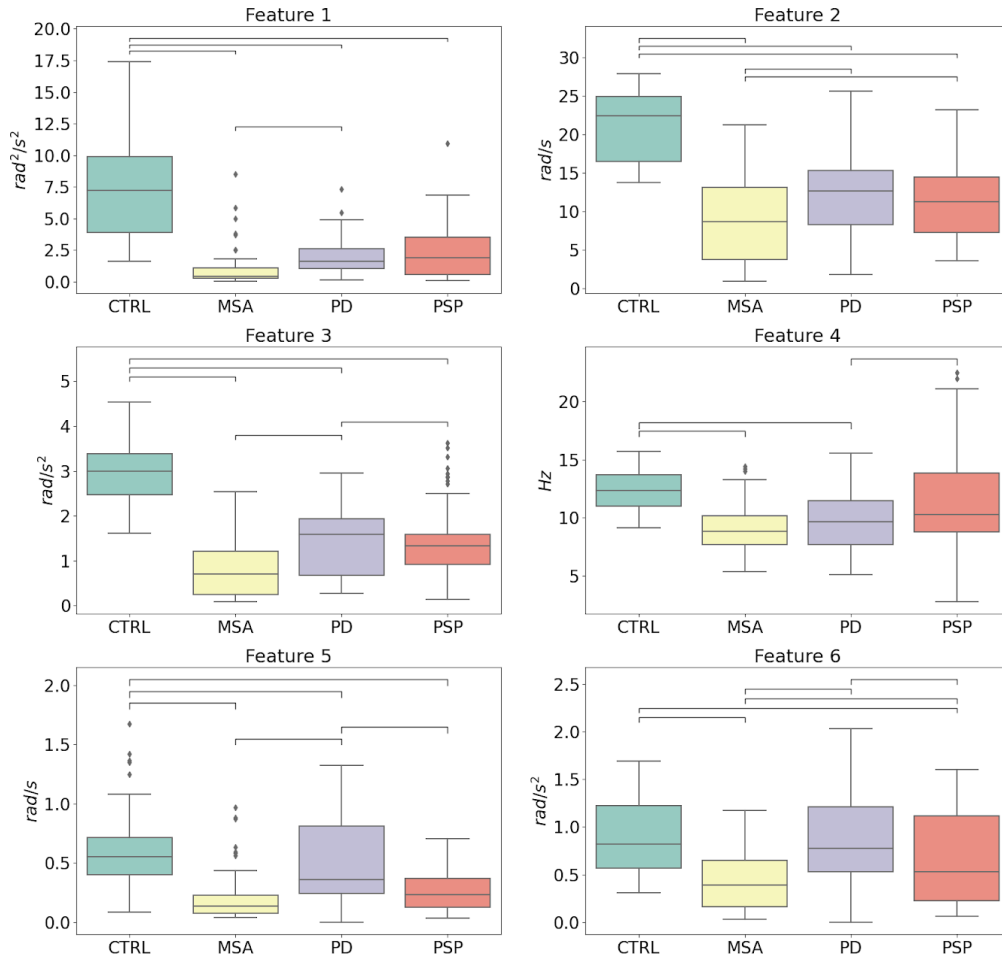
Figure 31 Distribution of selected features over participant groups. The lines above the boxplots denote significant differences between corresponding groups

Accuracy, precision and recall of the model trained on the selected features are given in Table 9. Figure 32 shows the confusion matrix that resulted from cross validation on each separate recording (i.e. prior to plurality vote on each patient's multiple recordings). The values and the color coding represent the number of cases in a corresponding category scaled to the total number of true cases for the class in question (row sums). The values in brackets show absolute numbers of cases.

The model assigned the correct diagnosis in a multiclass setting with overall accuracy of 76.11% of cases when inspecting all recordings individually. Out of all control samples, 94.23% were correctly classified, while out of all samples classified as controls, 76.56% were truly controls. The greatest confusion was seen between MSA and PSP groups, where 17.65% of those classified as MSA in fact belonged to the PSP group, and inversely 10.13% of those classified as PSP were diagnosed with MSA.

61

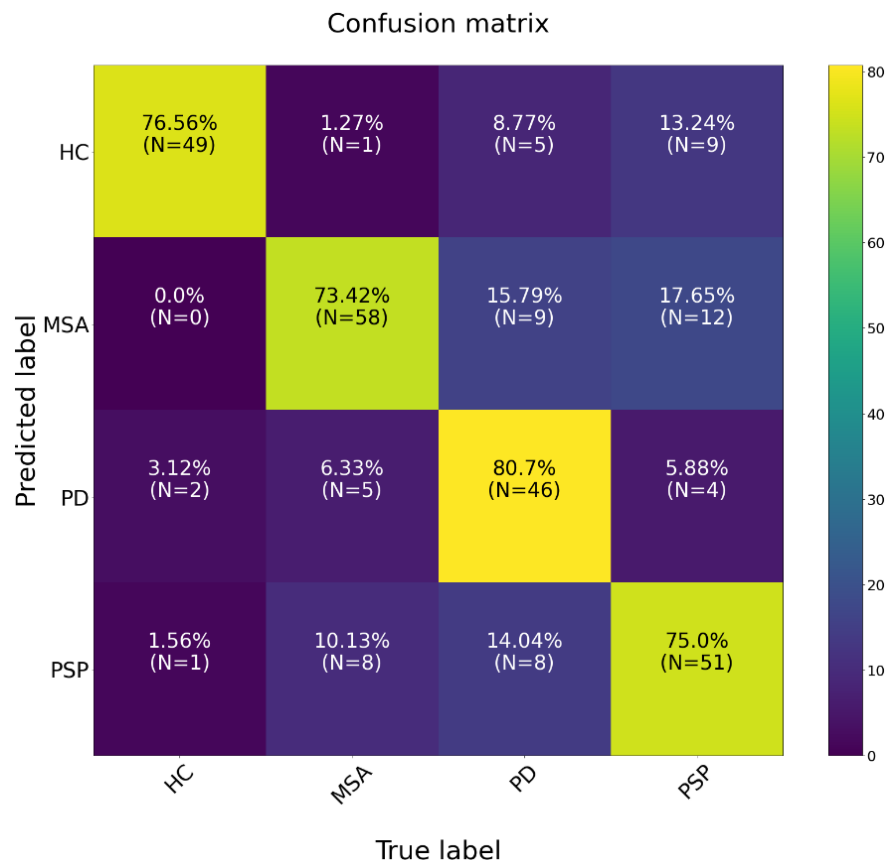|	Precision [%]	Recall [%]
HC	76.56	94.23
MSA	73.42	80.50
PD	80.70	67.65
PSP	75.00	67.10
Total accuracy		76.11%

Figure 32 Confusion matrix when observed for each single recording separately. The values on the main diagonal correspond to recall for the associated class.

When only the final predicted diagnosis for each patient is taken into consideration (final diagnosis being the most frequent one among the corresponding person's recordings), the overall accuracy becomes 85.18%. All 13 MSA patients were predicted to have the correct diagnosis, whereas 2 out of 16 PSP patients were incorrectly assigned to the MSA diagnosis. All healthy controls were properly recognized as well, with 3 out of 16 PSP patients falsely classified as healthy. Out of 14 PD patients, 11 were predicted to have the PD diagnosis, while 3 were wrongly found to carry the PSP diagnosis. In the set of patients that the algorithm classified as the PD group, all patients were clinically diagnosed with PD. Accuracy, precision and recall are given in percentages in Table 10. PSP was the most difficult diagnosis for the model to deal with, with recall of 68.75%, while MSA and HC groups were significantly different from the others, as reflected in the fact that 100% of them were properly classified.

Table 10 Precision and recall obtained on the test set for each class based on a single diagnosis per patient

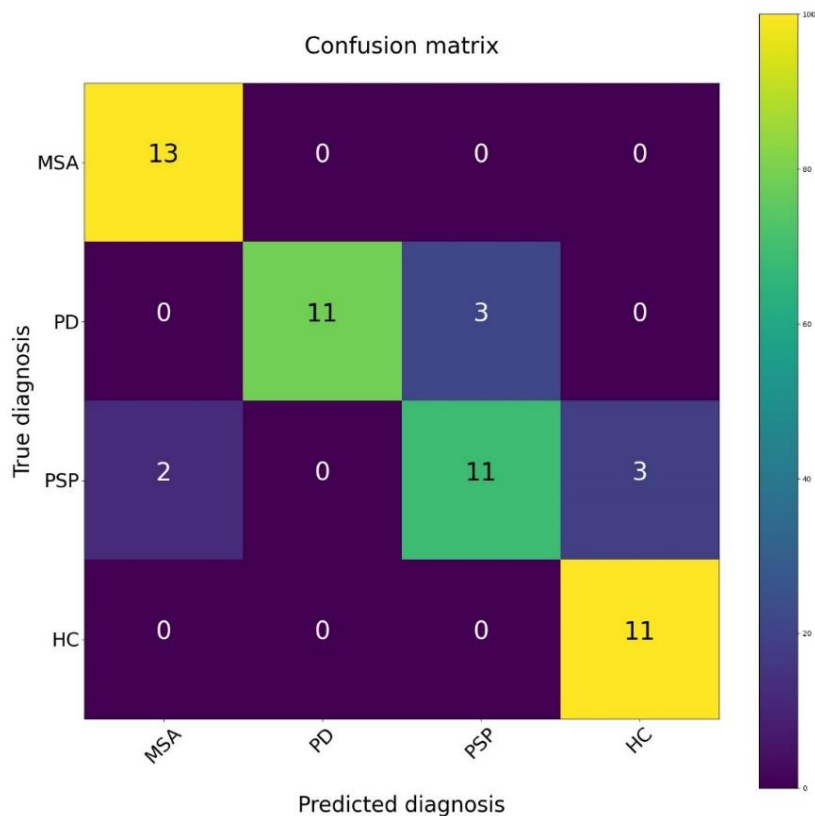|	Precision [%]	Recall [%]
MSA	86.67	100
PD	100	78.57
PSP	78.57	68.75
HC	78.57	100
Accuracy [%]		**85.18**

Figure 33 Confusion matrix expressed in absolute numbers of participants, when a single diagnosis was considered per patient

4.6.4. Discussion

After a preliminary test of several traditional machine learning algorithms, a kNN classifier was used to detect the presence of a neurodegenerative disorder of parkinsonian type and to output the most likely diagnosis based on gyroscope recordings of the thumb and the index finger during repetitive finger tapping. The algorithm was able to distinguish controls and three diagnoses on the patient level in 76.11% of cases on the level of single recordings, and 85.18% on patient level, demonstrating the potential of the used IMU-based system to contribute to the differential diagnostics of Parkinson's disease and atypical parkinsonisms. When one patient per diagnosis was considered, the breakdown of the results per group showed that MSA and HC diagnoses were most easily discerned (100%), whereas PSP was the hardest one (68.75%), due to some of these participants being incorrectly classified as MSA or HC groups.

The set of features that was found to be best able to tell apart the tested subject groups consisted of two features extracted from the gyroscope mounted on the thumb, and four features from the gyroscope on the index finger. Two of the features came from the most dominant finger opening axis (the rotation around the y-axis has the largest angles during tapping), concretely the root mean square of the index finger angular velocity and angular acceleration. One feature pertained to the mean left and right rotation of the index finger, i.e., how much the index finger leaned to the side while tapping, and one feature - spectral centroid of the vector of the index finger angular velocity - described the frequency content of the signal. This feature is determined by the tapping cadence but will also be influenced by other frequencies present in the signal, that could move the centroid up or down. For instance, the cadence of two recordings can be the same, but if higher frequencies are present in one of them, as would be the case if tremor is present, then this will be reflected in a greater centroid value. The features taken from the thumb gyroscopes included maximal rotation to the left or right, and variability of the maximal angular acceleration in each individual tap, taken as the standard deviation. The appearance of left-right rotation in the features is suggestive of pronation-supination hand tremor that is in PD described as "pill-rolling" [1].

Tests with deep learning approaches showed that the available tapping data was insufficient for these methods to demonstrate their potential. Collecting more data would make a better substrate for utilizing these approaches, where the system could leverage the pattern seeking abilities of deep learning rather than hand crafting the features. The scientific community would reap benefits from joint efforts to collect and analyze motion data in neurodegenerative disorders using a system like the one described, not only to improve the predictive power of machine learning approaches, but also to potentially use it to derive nuanced biomarkers that may have evaded our knowledge so far. We have seen the benefits of global collaboration on imaging, genomic and clinical data collected through the Parkinson's Progression Markers Initiative [181], which involves a large open database that has produced many scientific papers.

The system used here for recording the motion during finger tapping based on miniature gyroscope sensors is low cost, lightweight and easy to mount on a patient's hand. Though the system we used originally consisted of both accelerometers and gyroscopes, using just one modality for our analysis further reduced the cost of the system by reducing the hardware requirements. It also limited the system's storage demands and processing power requirements. Gyroscope recordings were chosen rather than accelerometer since the movements of interest during tapping were of rotational nature. It cannot be disregarded, however, that perhaps the addition of recordings from accelerometers and force sensors would further increase the predictive capability of the classification model.

So far, there have been no studies utilizing kinematic analysis of repetitive finger tapping for discerning Parkinson's disease and atypical parkinsonisms, although some studies have successfully used some form of finger tapping to detect the presence of Parkinson's disease against a group of healthy

participants [84], [182]. Keyboard typing has been used in a number of studies as a proxy for performing this test and collecting data for artificial intelligence algorithms to detect Parkinson's disease [77]–[79], [182]. The approach to diagnostic aid that was used in our work has the advantage of relying on a standardized clinical test, which may give it a head start regarding its adoption in clinical settings. Other modalities have been used to try and discern Parkinson's disease and atypical parkinsonisms, such as gait [183], [184] and speech [185], [186], and in the future perhaps a combination into a multi-modal approach would outperform single modality solutions.

It is of note that the groups of patients who participated in this study were a bit heterogeneous concerning their disability levels (mean Hoehn & Yahr group scores between 2.2 and 3.8). The best use of the system would be to recruit as many early-stage participants as possible, as this is the time when clinical diagnosis is the most elusive. This would however require a more complex, sort of a longitudinal study, where de-novo patients would be tested and recorded using the proposed system but would then go on to be followed for a number of years to make sure the correct diagnosis is reached with enough certainty to be useful for training the computer assistant and its machine learning algorithm. Although the diagnoses in our data set were carefully determined by clinical experts of clinical neurology, they were not confirmed at autopsy, and this may have affected the results. Given the imperfection of the reference diagnoses, another approach to the analysis would be to treat it as a weakly supervised problem [187], or even to develop algorithms which would draw conclusions in an unsupervised manner, but this would again require a larger amount of data than was collected for this study.

Future work should include collection of additional recordings from different patients, particularly in the early stages of the disease.Solutions for damping the effect of possible noise in the labels (non-autopsy-confirmed diagnoses) is also a point to be addressed, and so is turning the presented analysis into a user-friendly application that could be adopted in the clinical settings, through working closely with experts of clinical neurology.

Proposed technology in light of recent events

Healthcare systems and health services were put under tremendous stress by the Covid-19 pandemic that started early in 2020 and in the moment of writing in 2022 has somewhat subsided but is still very much affecting the processes, making already complex systems and procedures even more complicated. Frequent rotation of the medical staff was required, breaking continuity in care provided to patients by their assigned doctors, elective procedures were put on hold, non-critical patients were dedicated a minimal amount of time, some were rerouted to different institutions, and others were told to wait. Some other new infections have also recently disturbed the public, such as the appearance of monkey pox, and certain pharmaceutical companies have started working on producing a vaccine for the nipah virus, as there has been some suggestion that nipah is what may bring on the next great pandemic. It is much deadlier than Covid, the spread of which may be taken as a drill for potentially something more dangerous to come. Either way, it seems that some measures of precaution and changes in operation of healthcare institutions are here to stay. Technology has recognized this to a large extent, putting more effort into coming up with solutions for telemedicine, contactless sensors, computer assistants, and more. In June 2022, papers wrote about the amount of seed stage funding in Europe that was brought in by industry and reported that health tech was leading the charts (118 million euros), surpassing fintech (77 million euros) which used to dominate the seed funding landscape up until this point [188]. In that light, a system such as the one described here could be a timely assistance to help take some of the burden off the medical staff and speed up the diagnostic process. It could enable interdisciplinary teams of engineers and clinical doctors to apply the same or similar system for diagnostic aid and share recordings and

possibly draw deeper insights into the subtle differences between the pathologies, made possible by polling of multi-center open data. A review on submitted papers focusing on machine learning applications in health research emphasized the importance of well-annotated data that is easily accessed, combined with increased involvement of clinical staff in the application development process [189].

5. ANALYSIS OF GAIT IN PARKINSON'S DISEASE IN DUAL-TASK PARADIGM

This chapter aims to use gait kinematic data and artificial intelligence to answer the questions posed in Hypothesis 2 and Hypothesis 4 from a different angle. **Hypothesis 2** assumes that with the help of artificial intelligence, patients with PD can be discerned on the individual level from persons without neurological disorders, and **Hypothesis 4** states that is possible to programatically choose a subset of relevant features extracted from kinematic signals which increase the performance of classification among the observed disorders, or in this case - between healthy controls and patients with Parkinson's disease.

We will use methods of artificial intelligence to identify the subset of spatio-temporal parameters which are the most useful for discerning *de novo* PD patients from participants with no neurological disorders and build on top of these parameters a classification model as a means of assistance in PD diagnostics.

Gait analyses will not be used to test Hypothesis 1 and Hypothesis 3, as they pertain to multi-group differences and involve patients suffering from atypical parkinsonisms, for which no gait data is available at this time. This section describes kinematic data describing gait in healthy controls and participants with PD.

5.1. Participants

Forty *de novo*, drug naive patients were consecutively recruited at the Neurology Clinic, Clinical Centre of Serbia, Belgrade. Inclusion criteria included Step I and Step II criteria of the UK Parkinson's Disease Society Brain Bank (PDSBB) Diagnostic Criteria [43][6]. The study included participants with unilateral presentation of disease with or without axial involvement (H&Y stages 1 or 1.5), without regard for disease duration. MDS UPDRS Part III [33] and H&Y [48] staging systems were used for clinical evaluation, and mental status was rated through the Hamilton Depression Rating Scale (HAM-D) [190], Beck Depression Inventory (BDI) [191], Hamilton Anxiety Rating Scale (HAM-A) [192] and the Apathy Scale (AS) [193]. Cognitive screening was done using the Mini-Mental State Examination [194], and more deeply assessed, using the Addenbrooke's Cognitive Examination Revised (ACE-R) [195]. The control group was composed of forty healthy, age and gender balanced participants (Table 11). Exclusion criteria included any condition that could interfere with motor activity, be it of neurological, orthopedic, or other medical nature. The study was performed in accordance with the ethical standards of the Declaration of Helsinki and its later amendments. Medical Ethics Committee of Clinical Centre of Serbia approved the research protocol, and written informed consent was obtained from each participant.

Table 11 Demographic and clinical data for the tested participants

	PD patients (n=40)	Controls (n=40)	P
Age (years)	59.83±10.57	59.79±11.85	0.989
Sex (females)	16(40%)	19(47.5%)	0.712
Sex (males)	24(60%)	21(52.5%)	
Education (years)	13.63±2.71	12.98±2.96	0.302
Disease duration (years)	1.38±1.16	/ /	/
MDS-UPDRS Part III	15.8±5.49	/ /	/
Gait item from MDS-UPDRS	0.68±0.53	/ /	/
H&Y	1.16±0.24	/ /	/
HAM-A	4.45±5.12	5.79±5.71	0.264
HAM-D	5.23±5.92	2.60±3.28	0.016*
BDI	7.43±6.96	3.79±4.02	0.005*
AS	10.65±7.17	5.60±6.02	0.001*
MMSE	28.7±1.20	29.47±0.98	0.476
ACE-R	91.6±5.47	91.95±6.90	0.737

Values are given as mean ± standard deviation; MDS-UPDRS- Movement Disorder Society Unified Parkinson's Disease Rating Scale; H&Y - Hoehn and Yahr Staging system; HAM-A - Hamilton Anxiety Rating Scale; HAM-D - Hamilton Depression Rating Scale; BDI - Beck Depression Inventory; AS - Apathy Scale; MMSE - Mini Mental State Examination; ACE-R - Addenbrooke's Cognitive Examination Revised

---

[6] Parts of the analysis presented in this chapter have been adapted and published in [113].

5.2. Instrumentation and protocol

Instrumentation used in this chapter involved the GAITRite portable electronic walkway (CIR Systems, Haverton, PA), mat of 5.5m active area with built-in pressure sensors with 1.27cm spatial resolution and maximum sampling rate of 240Hz. The walkway is intended to provide spatial and temporal information on the subject's gait. The dedicated system software uses physics and mathematics to calculate gait parameters (e.g., velocity). Participants were recorded walking up and down the corridor in comfortable shoes for 6 passes in each of the following three scenarios: (1) basic walking, pacing at the subject's comfortable speed, (2) motor task, during which the subjects walked on the mat while holding a glass of water and trying not to spill it, and (3) mental task, during which the participants were asked to recursively subtract 7 starting from 100 [196]. Each task amounted to roughly 50m of walking.

5.3. Gait data

The data used for analysis comprised the 37 parameters obtained from the GAITRite walkway system, grouped by participant and task, and averaged over 6 recorded passes. The initial parameters are listed in Table 12. Three tasks with 37 parameters resulted in 111 variables total per participant.

Table 12 List of initial spatio-temporal gait parameters obtained from the GAITRite system

1	Velocity	20	Heel Off On Time
2	Normalized Velocity	21	Heel Off On %
3	Cadence	22	Double Support Load Time
4	*Step Time Differential	23	Double Support Load % of Cycle
5	Step Length Differential	24	Double Support Unload Time
6	*Stride Time Differential	25	Double Support Unload of Cycle
7	Step Time	26	Stride Velocity
8	Step Length	27	Step Length CV
9	#Stride Time	28	*Step Time CV
10	#Stride Length	29	#*Stride Length CV
11	Heel To Heel Base Support	30	#Stride Time CV
12	#Swing Time	31	#Swing Time CV
13	Stance Time	32	Stance Time CV
14	*Single Support Time	33	*Stride Velocity CV
15	#Double Support Time	34	*Single Support Time CV
16	Swing % of Cycle	35	#Double Support Time CV
17	Stance % of Cycle	36	Heel Off On CV
18	Single Support Time % of Cycle	37	*Heel-to-Heel Support Base CV
19	Double Support Time % of Cycle		

***** CV –Coefficient of variation

Normalized velocity refers to velocity divided by the average leg length. Heel off/on time represents the time between heel-off and heel-on points of two consecutive steps made by the same foot. Heel-to-heel base support illustrates the perpendicular distance between the center of the heel on one foot and the line of progression that the opposite foot forms. Step and stride time differential illustrate asymmetry through the difference seen in the step and stride times respectively between the left and the right foot. Step length differential represents the difference in step lengths between the left and the right foot.

5.4. Classification: Parkinson's disease vs healthy controls

Given that high data dimensionality, as reflected in 111 parameters per subject, often negatively impacts classification algorithms, and also makes model interpretation more difficult [197], we sought to reduce the number of parameters that would be used to detect the presence of PD. We started with pair-wise cross-correlating all variables using Pearson's correlation, with the aim of excluding highly correlated variables. Since simply ranking the correlation coefficients and eliminating the highest ones would bring into question which of the highly correlated features should we keep, an alternative approach was employed. Namely, the pair-wise correlation was taken to be the similarity metric between two variables, and using this similarity as input, the features were clustered via affinity propagation. The exemplars of the obtained clusters were then passed to the next step of the algorithm, which involved ranking the features by importance using random forests [198], and relying on mean decrease in accuracy which illustrates how much the model accuracy suffers when out-of-bag data for the variables are randomly permuted. The higher the accuracy drop for permuting a particular variable, the more important that variable is in discerning PD from healthy participants. Highest ranked features were then used as input to the classification step. The flowchart of the processing pipeline is given in Figure 34. The reason for including the correlation-based clustering step prior to random forest is that random forest feature selection method is known to suffer from correlation bias [199], i.e., correlated groups of features lose in their assigned weights as the size of the mutually correlated group grows, so the importance attributed to these variables does not reflect their true value to the model.

Affinity propagation is a clustering method focused on finding a subset of representative points, exemplars. In the famous k-means clustering this would be an analogue of cluster centroids, although a centroid can be a point that does not exist in the dataset, but an exemplar is a real data point. K-means starts by randomly choosing the centroids and then iteratively improving the split, whereas affinity propagation takes into consideration the pair-wise similarity metrics given at input [200]. And while k-means requires the number of desired clusters to be specified in advance, affinity propagation does not. This may mean that the clustering in k-means will be dependent on initial conditions, while affinity propagation tries to remedy this. Affinity propagation does, however, require other input parameters, namely "preferences", $s(k,k)$. These are real numbers given for each data point, so that larger values make the datapoints more likely to be chosen as exemplars. These preferences can be set to a common value to indicate that all datapoints are equally valid choices of exemplars, as was done in this work. Affinity propagation observes all data points simultaneously and regards them to be nodes in a network, the edges of which exchange real-valued messages. The messages are recursively updated, and their magnitude describes the affinity of data points to one another. There are two kinds of messages exchanged, "responsibility" $r(i,k)$ and "availability" $a(i,k)$, the first one representing how well point k is suited to be the exemplar for point i, and the second one tells the point i how appropriate it would be for it to take point k as its exemplar. Responsibilities and availabilities are updated as:

$$r(i,k) \coloneqq s(i,k) - \max_{k`\neq k}\{a(i,k`) + s(i,k`)\} \tag{17}$$

$$a(k,i) \coloneqq \min\{0, r(k,k) + \sum_{i`\neq i, i`\neq k} \max\{0, r(i`,k)\}\} \tag{18}$$

With self-availability a(k,k) updated as:

$$a(k,k) \coloneqq \sum_{i`\neq k} \max\{0, r(i`,k)\} \tag{19}$$

These messages are combined to decide the exemplars and clusters. Implementation of affinity propagation in R package 'apcluster' was used in this study.

Random forest is a machine learning model that combines many weak predictors – decision trees – to make an ensemble strong predictor. Each tree node tests the input data against a threshold and determines if it passes or not. The subregion created after such a split can then again be divided based on some threshold-based criterion. Each decision tree's prediction is considered a vote and the one that takes the majority (or plurality) vote wins. The trees are constructed on a random subset sampled with replacement from the set of learning samples, but also from the set of features. This property allows the assessment of importance of variables for the given model [201], and this approach has been used in various fields, such as psychology [202] or bioinformatics, for instance to decide on a subset of genetic markers that contribute most to a certain disease [203]–[205]. Consider there are N input variables and m data points $x_1, x, ..., x_m$. Having constructed the forest, the out-of-bag values of the k-th variable are randomly permuted and such a permuted dataset is used for inference in the corresponding trees, and the predictions are saved. This is repeated for $k = 1, 2, ..., N$. The plurality vote is then found for each x_i and compared to the true class label to obtain a misclassification rate. The measurement of interest that is then used to sort variables is the increase in misclassification rate when the data are permuted compared to when they are left as is [201].

Cross- Correlation   Affinity propagation clustering   Random forest feature importance   Classification

GaitRite data acquisition
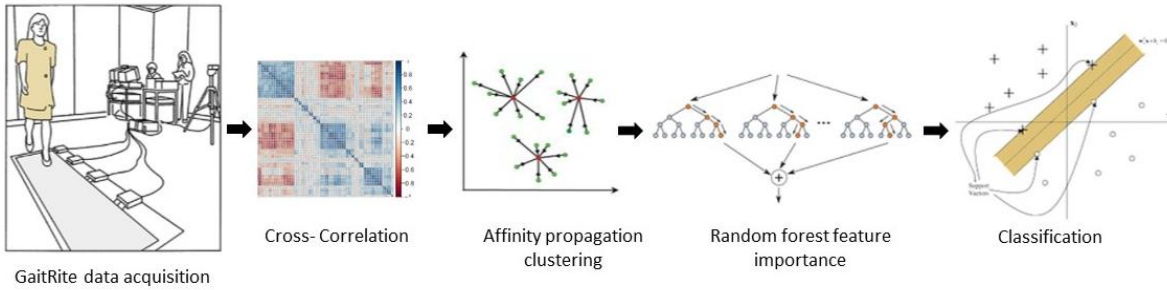
Figure 34 High level flowchart of the performed analyses (Adapted from [113]*)*

A support vector machine (SVM) model (see Chapter 4.5.1.) was built on input data with selected variables to discern healthy participants from PD patients. A parallel SVM model was build on the same dataset but using a different set of variables, namely 8 parameters commonly used in gait analysis: mean values for stride time, stride length, swing time, double support time and their corresponding standard deviations [24], [196], [206]. The results of the two models were compared for the three test tasks: base, motor and mental, as well as for their combination. Classification performance was assessed using 10-fold cross validation. Metrics of interest were sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).

$$Sensitivity = \frac{No\ of\ true\ positives}{No\ of\ actual\ positives} * 100\% \tag{20}$$

$$Specificity = \frac{No\ of\ true\ negatives}{No\ of\ actual\ negatives} * 100\% \tag{21}$$

$$PPV = \frac{No\ of\ true\ positives}{No\ of\ all\ points\ classified\ as\ positives} * 100\% \tag{22}$$

$$NPV = \frac{No\ of\ true\ negatives}{No\ of\ all\ points\ classified\ as\ negatives} * 100\% \tag{23}$$

Variable importance was also assessed on a 10-fold basis.

The analyses and plotting were performed using R v. 3.2.2.

5.5. Results

Pair-wise cross-correlation showed certain groups of variables were highly correlated with each other, as illustrated in Figure 35. For instance, double support time was found to be highly negatively correlated with cadence, with correlation coefficient of -0.876. Correlation persisted through all three tasks, as well as when all tasks were analyzed combined (Figure 36).
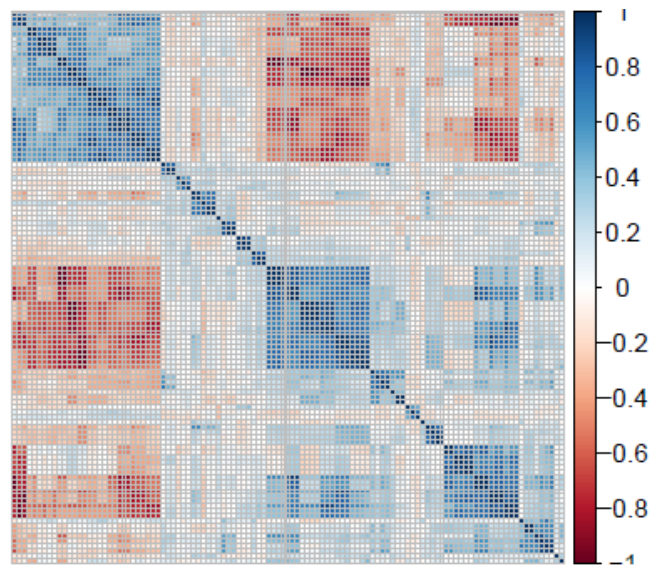
Figure 35 Correlation matrix shows high correlation between some groups of variables
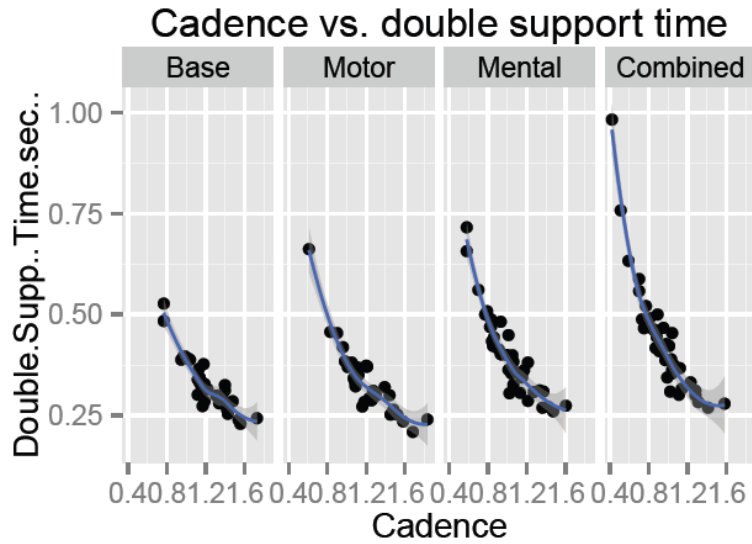
Figure 36 High negative correlation between cadence and double support time

When pair-wise cross-correlation was passed to affinity propagation clustering as the similarity metric, 19 clusters were obtained, and their exemplars were kept for further processing (Table 13). This reduced the number of features from the initial 37 to 19.

Table 13 Parameters kept for further processing after clustering

Normalized velocity	Step length CV
Step time differential	Step time CV
Step length differential	Stride length CV
Stride time differential	Stance time CV
Stride time	Stride velocity CV
Stride length	Single support time CV
Single support time	Double support time CV
Double support time	Heel off on CV
Heel off on time	Heel-to-heel support base CV
Double support load % of cycle	

The 19 exemplars established by affinity propagation were passed to a random forest algorithm for feature ranking, which yielded the final eight most relevant parameters: stride length and its coefficient of variation (CV), stride time and stride time CV, swing time and swing time CV, step type asymmetry and the CV of heel-to-heel base support. The boxplots of these eight parameters are shown in Figure 38, split by group and by task type, and a corresponding table is given containing mean values and inter-quartile range (Table 14). The importance ranking showed however that one and the same parameter may be more prominent in one task over another. Stride time CV was ranked highly for motor and mental tasks, but not for the base task. We can see from the boxplots that the values for stride time CV are largely overlapping between groups for the base task, as opposed to the dual tasks. Swing time CV in the base task was ranked relatively low, while in the mental task it was found to be among the 5 most

73

important features. Stride length CV during performance of the mental task was found to be the most important variable that stood out among all others. Figure 37 shows the 10 most important features as ranked by the random forest algorithm.

Table 14 Selected parameters grouped by task and test group, given as median (interquartile range)

Parameter	Base Task		Motor Task		Mental Task	
	HC	PD	HC	PD	HC	PD
Step Time Asymmetry [s]	0.008 (0.004-0.013)	0.018 (0.007-0.027)	0.008 (0.003-0.012)	0.018 (0.006-0.033)	0.012 (0.005-0.017)	0.023 (0.009-0.037)
Stride Time [s]	1.064 (1.017-1.105)	1.091 (1.058-1.184)	1.056 (1.01-1.116)	1.110 (1.055-1.183)	1.129 (1.075-1.202)	1.215 (1.135-1.304)
Stride Time CV	0.026 (0.021-0.035)	0.029 (0.024-0.037)	0.023 (0.02-0.028)	0.031 (0.023-0.035)	0.032 (0.025-0.042)	0.051 (0.035-0.084)
Stride Length [cm]	129.39 (115.74-137.85)	122.12 (112.56-132.53)	122.29 (112.35-132.38)	119.47 (107.69-125.31)	119.78 (108.96-131.85)	111.344 (100.35-121.89)
Stride Length CV	0.026 (0.021-0.036)	0.035 (0.028-0.043)	0.026 (0.021-0.027)	0.035 (0.027-0.043)	0.039 (0.03-0.048)	0.058 (0.042-0.082)
Swing Time [s]	0.377 (0.359-0.396)	0.393 (0.378-0.419)	0.368 (0.356-0.394)	0.394 (0.375-0.411)	0.394 (0.369-0.411)	0.416 (0.392-0.444)
Swing Time CV	0.058 (0.046-0.08)	0.045 (0.037-0.058)	0.052 (0.044-0.064)	0.046 (0.038-0.055)	0.064 (0.054-0.136)	0.066 (0.052-0.082)
H-H Base Support CV	0.183 (0.15-0.227)	0.163 (0.143-0.182)	0.201 (0.161-0.244)	0.17 (0.148-0.196)	0.203 (0.17-0.23)	0.155 (0.135-0.189)

*CV – coefficient of variation; HC – healthy controls; PD – Parkinson's Disease

The chosen set of parameters differs from the standard set found in literature in that it does not feature double support time and its CV, but instead contains step time asymmetry and CV of heel-to-heel base of support. Our data did not show significant between group differences in double support time and its variation, as shown in Figure 39, and variable importance ranking suggested that the inclusion of these parameters would in fact decrease the classification accuracy.

When the obtained 8 features were fed to an SVM classifier combining all performed tasks, the mean classification accuracy was 85%, sensitivity and specificity both 85%, PPV 86% and NPV 89%, meaning the predictions of the classifier could be roughly equally trusted for the PD and the control groups. The SVM model performed worse when it was fed with the standard set of parameters, achieving 80% in accuracy, sensitivity and specificity, 82% PPV and 78% NPV (Table 15).

Classification was also performed when the input parameters were selected for one single task at a time (base, motor, or mental), rather than their combination. The combination of all tasks had the best performance both for the standard features and those selected in our work. The selected features showed a decrease in classification performance in the base task compared to the motor and mental tasks, while this was not the case for the standard set of parameters, where the base task even had a slight advantage over the others.
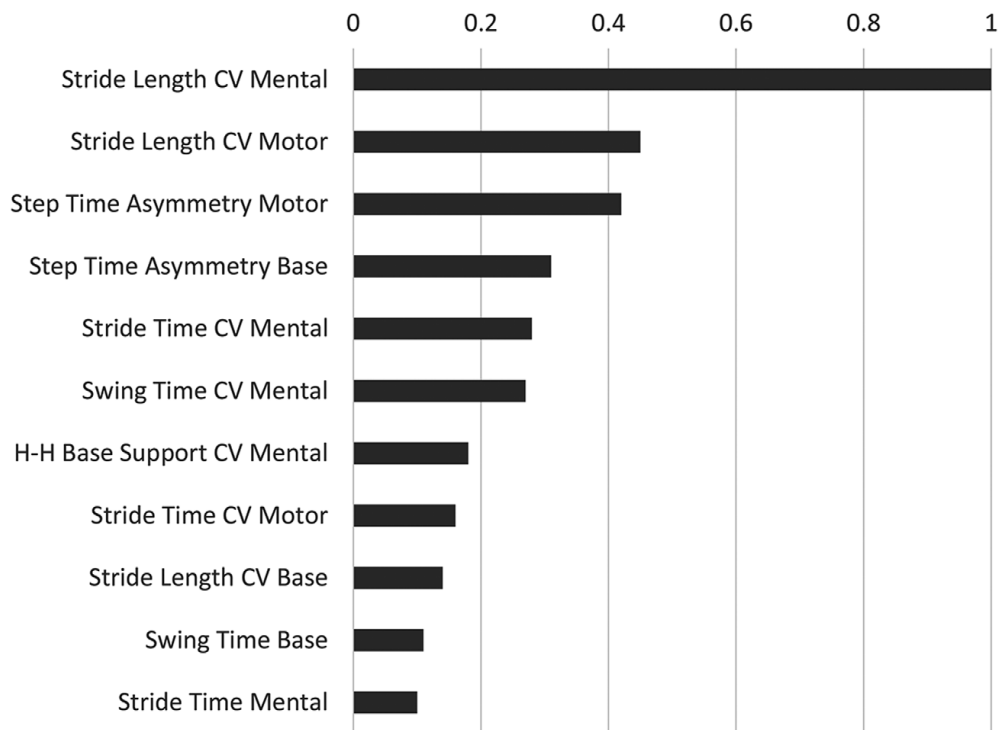
Figure 37 Permutation-based variable importance normalized to 1. The 10 highest ranked parameters are shown (Printed in [113]*).*
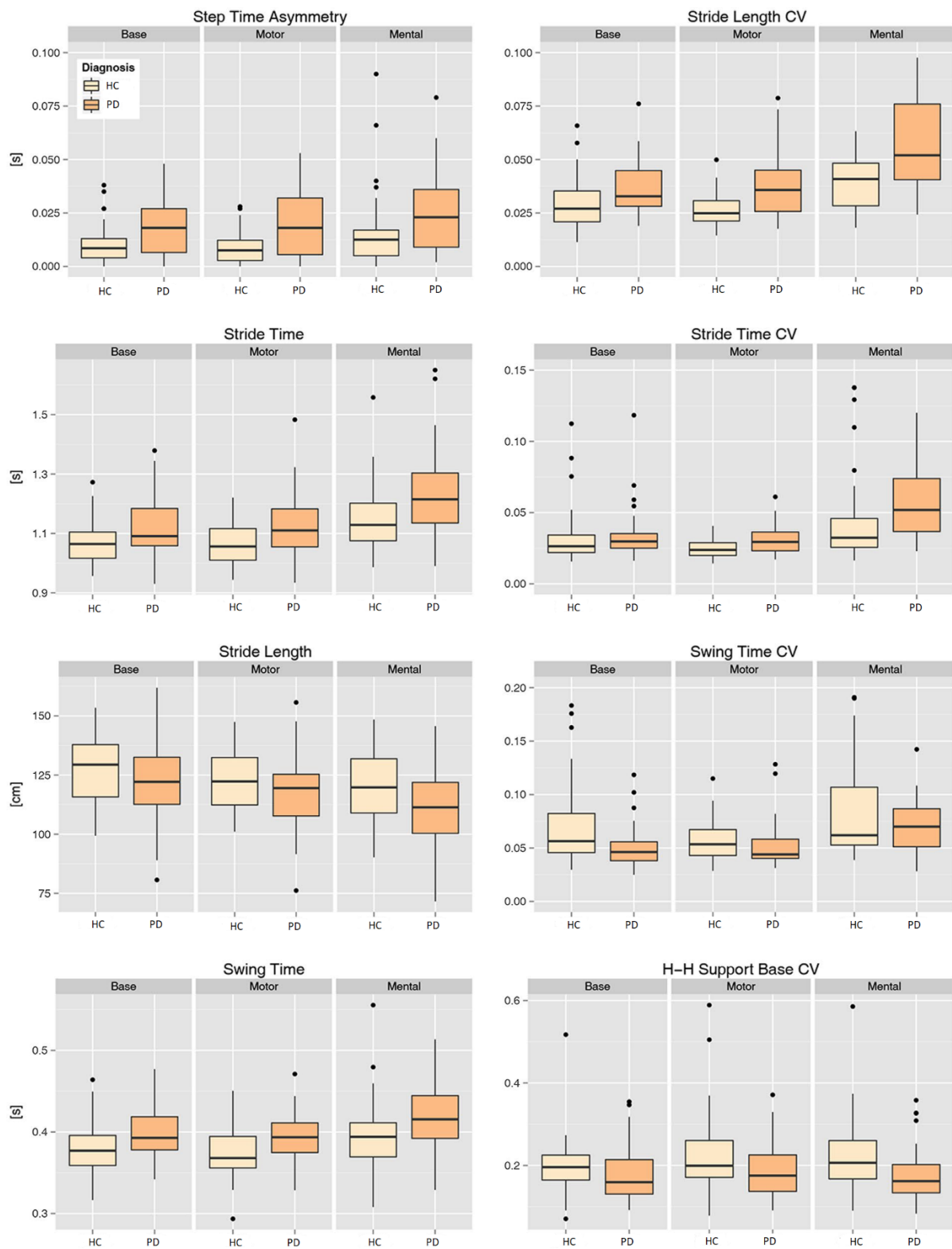
Figure 38 Selected gait parameters used to build classification models, shown as boxplots for healthy controls and PD patients, in three tested gait scenarios: base, motor and mental tasks (Adapted from [113]*)*

Table 15 Classification performance obtained for the standard and the suggested set of parameters in base, motor, mental and combined tasks

	Standard set of parameters				Suggested set of parameters			
	All 3 tasks	**Base**	**Motor**	**Mental**	**All 3 tasks**	**Base**	**Motor**	**Mental**
Accuracy	**0.80±0.11**	0.75±0.16	0.74±0.12	0.74±0.12	**0.85±0.10**	0.75±0.14	0.79±0.14	0.82±0.13
Range	0.62-1.00	0.50-1.00	0.50-0.87	0.62-0.87	0.75-1.00	0.50-1.00	0.62-1.00	0.62-1.00
Sensitivity	**0.80±0.16**	**0.90±0.13**	0.75±0.17	0.60±0.24	**0.85±0.13**	0.77±0.18	0.80±0.20	0.82±0.21
Range	0.50-1.00	0.75-1.00	0.50-1.00	0.25-1.00	0.75-1.00	0.50-1.00	0.50-1.00	0.50-1.00
Specificity	**0.80±0.20**	0.60±0.21	0.72±0.22	**0.87±0.13**	**0.85±0.13**	0.72±0.22	0.77±0.18	0.82±0.17
Range	0.50-1.00	0.25-1.00	0.50-1.00	0.75-1.00	0.75-1.00	0.25-1.00	0.50-1.00	0.50-1.00
PPV	**0.82±0.13**	**0.85±0.20**	0.77±0.16	0.71±0.14	**0.86±0.12**	0.77±0.18	0.81±0.17	0.85±0.16
Range	0.60-1.00	0.50-1.00	0.50-1.00	0.57-1.00	0.75-1.00	0.50-1.00	0.60-1.00	0.60-1.00
NPV	**0.74±0.21**	0.70±0.20	0.65±0.11	**0.78±0.22**	**0.89±0.16**	0.72±0.18	0.73±0.15	0.82±0.17
Range	0.40-1.00	0.40-1.00	0.50-0.80	0.40-1.00	0.57-1.00	0.50-1.00	0.57-1.00	0.60-1.00

* PPV – Positive Predictive Value, NPV – Negative Predictive Value
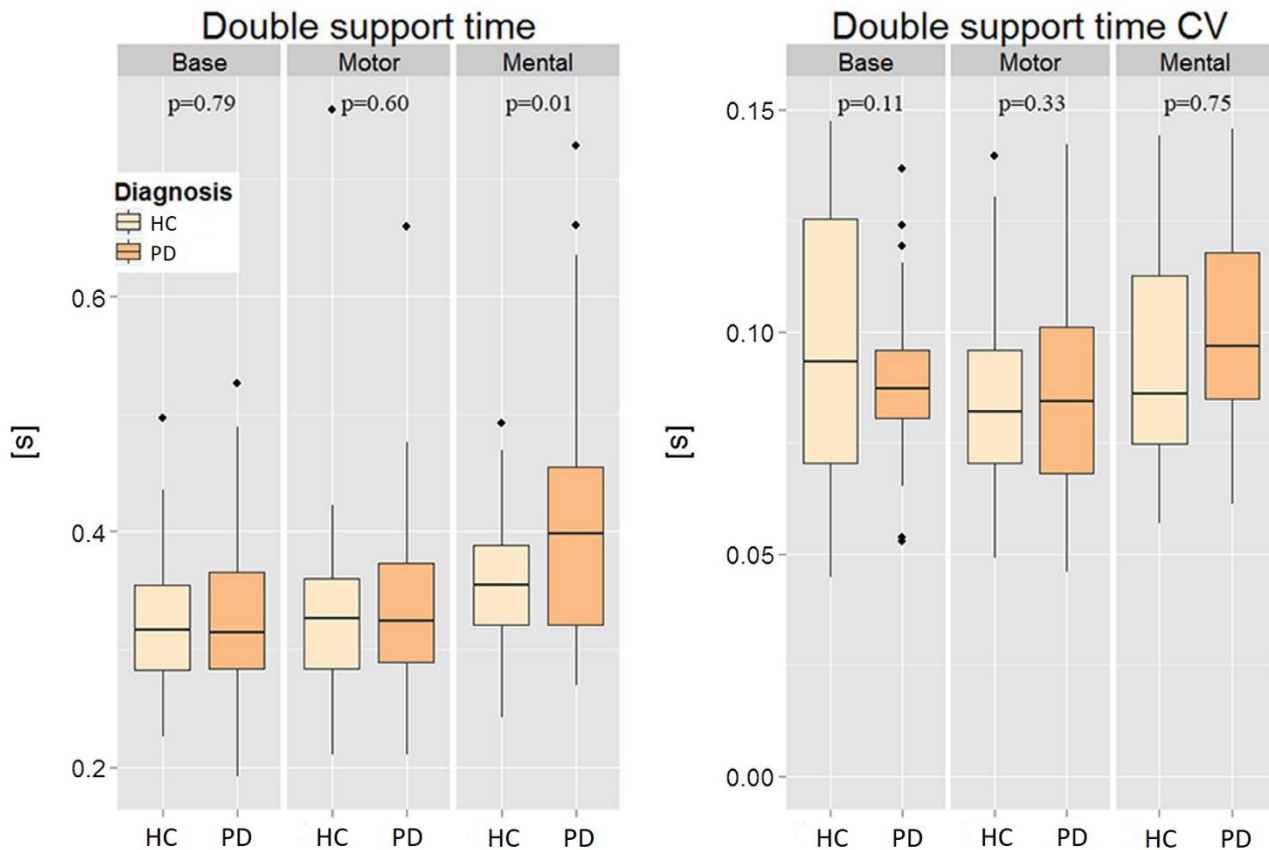
Figure 39 Boxplots for double support time and its coefficient of variation (CV). P stands for P-value obtained using unpaired t-test (Adapted from [113]).

5.6. Discussion

In this chapter we analyzed data on gait collected using an electronic walkway which provides a plethora of spatio-temporal features of gait, however without clear standardization on feature selection for detection of PD pathology. The goal was to identify the parameters that contribute the most to differentiating early PD from healthy controls on the patient level, and subsequently utilize them as input to a classification algorithm that would suggest whether an individual is suffering from PD. The walkway and data manipulation allow for quantification of gait parameters whose subtle differences between healthy persons and those with a neurological disorder could be captured when they are not so easily observed by the naked eye.

The performed analysis yielded the following eight parameters as the most relevant for classification: stride length and stride length CV, stride time and stride time CV, swing time and swing time CV, step time asymmetry, and heel-to-heel base support CV. When these parameters were used as input to an SVM classifier, the achieved accuracy was 85%, which was higher than when the same classifier was used with input parameters which were handpicked to correspond to those commonly found in literature (78%), and which include: stride time, stride length, swing time and double support time, along with their CVs. Authors in [207] assessed the gait of PD in comparison to healthy individuals and found significant differences between the groups in parameters including stride length, cadence, stance time and double support time. Using these features, they built a classifier which provided 80.4% accuracy. The choice of parameters used in this analysis has a notable overlap with those commonly found in literature, confirming their relevance, albeit with some differences. While previous studies found double support time to be significantly different between PD and control groups, our findings failed to reproduce that finding, although there was a tendency of group means for double support time and its variance to differ if observing the mental task specifically, though the difference did not reach statistical significance. This may mean that double support time may be a more reliable predictor later in the disease progression than it is in early stages. As one of the highly relevant features this analysis introduces step time asymmetry, which stands for difference in step duration between the left and the right legs, as an indicator of asymmetric involvement in PD. Increased gait variability was found to be characteristic of PD, reflected in the observation that multiple selected parameters were coefficients of variation. This confirms the findings of earlier research work [24].

The feature ranking algorithm showed the single most important variable for detecting *de novo* PD from gait was stride length variation during the mental task, suggesting impairment of gait automation as an early sign of PD. Another relevant parameter that emerged in this analysis was the coefficient of variation of heel-to-heel base support, which was found to be smaller in PD than in healthy controls. Patients with PD are known to walk with a wider base of support than healthy persons possibly to counteract the fear of falling [208]. Although out data did not corroborate this finding, as we saw no significant differences in base of support, it is possible that the detected reduction in the variability of the support base has the same roots as a compensation strategy in the face of the risk of falls.

High predictive power of the performance on the mental task was seen in the results of training the classifier on data describing one gait task at a time, where the mental task only provided the highest results of all three (82% accuracy), followed closely by the motor task (81%) and the lowest for the base task (76%), although the combination of all three tasks outperformed each task taken individually (85%). This deterioration of performance when walking with a competing dual task reflect the difficulties that PD patients face in dividing attentional resources [22], and is in line with the findings of a meta-analysis which observed deterioration of gait in PD when a dual task is introduced, regardless of the baseline level, and irrespective of the type of the interference task, whether motor, arithmetic, or other [209]. This

superiority of the mental or motor task in detecting early PD however was not apparent when the classifier was trained using the standard hand-picked set of features. In this case all three tasks individually had a comparable accuracy (74-75%), while the combination of all three tasks gave the best results (80% accuracy).

The gait parameters identified as the most important in this study emphasize gait variability and left/right asymmetry, particularly in the mental task, reflecting difficulties in attention resource management in early PD. The variable ranking and task importance may change over the course of the disease, which could be tested through comparison with a later-stage group of patients that would be recruited and recorded in the same protocol. There are indications in literature that dual task walking may provide clues for differential diagnostics too. A study used gait analysis with a cognitive task for discerning PD from PSP patients and found that PSP gait was poorer than PD already in early stages, showing a greater reduction in gait speed, and increased cadence and length of cycle in comparison with PD [210].

These results support Hypothesis 4, showing that AI can be used to programmatically select parameters informative for PD diagnostics, and Hypothesis 2, demonstrating that AI can offer help in discerning patients with PD from neurologically healthy individuals. This group of tests did not involve patients with atypical parkinsonisms, so the hypotheses that pertain to differentiation of PD from atypical parkinsonisms were not tested using gait kinematics, but only through finger motion.

Future work should include recruiting patients with atypical parkinsonisms and using the described analysis of gait to test Hypothesis 1 and Hypothesis 4, i.e. whether statistical analyses and artificial intelligence can help guide differential diagnostics between PD and atypical parkinsonisms on group and patient level.

6. CONCLUSION

Clinical diagnosis of Parkinson's disease and atypical parkinsonisms remains a challenging and time-consuming task. Recent advances in neuroimaging have made this modality a powerful ally in clinical practice, however the price, required time, limited availability in hospitals and the use of radioactive agents in some cases do not make this a ubiquitous solution, and quicker and cheaper options are desired. Kinematic analyses provide a plausible alternative, particularly when combined with the perks brought in by algorithms of artificial intelligence and machine learning. We hypothesize that this synergy between motion sensors and artificial intelligence can positively contribute to improving the speed and accuracy of obtaining a correct diagnosis, which would in turn allow better treatment adjustment and facilitate participant selection in clinical trials, which would particularly be of interest in early stages of the disease.

As its main focus, this study sought to utilize artificial intelligence combined with kinematic finger tapping data obtained by a low-weight, low-cost inertial sensor setup. Repetitive finger tapping is a simple task, borrowed from the UPDRS battery of tests, and shown to correlate better with the overall UPDRS score than other forms of repetitive motions.

Three groups of patients have been recruited, namely individuals suffering from PD, PSP, and MSA, as well as a group of controls without neurological disorders. We tested whether certain parameters can be determined through such a system that would point to group level differences in finger tapping between the examined groups (Hypothesis 1), whether using the finger tapping signals AI can help detect PD and differentiate it from healthy controls (Hypothesis 2), if AI can help differentiate PD not only from controls, but also from atypical parkinsonisms (Hypothesis 3) and whether particular features can be extracted programmatically that would offer the most information useful for PD diagnostics (Hypothesis 4).

Statistical analysis of finger tapping parameters showed certain differences in trends between the tested groups, thus confirming Hypothesis 1 of this thesis. The most notable finding was the lack of progressive temporal reduction in tapping amplitude for the majority of PSP patients and HC, which was not the case for PD and MSA groups.

When time-frequency analysis was applied as a preprocessing step to the gyrposcope signals recorded from healthy individuals and PD patients, and used as input to a multi-layer perceptron neural network, the patient group was well discerned from the controls, with accuracy of 92%, showing that AI teamed with kinematic processing can indeed help discern PD patients from healthy controls, as posed in Hypothesis 2.

When other patient groups were included in the analysis, a kNN model was trained on programmatically selected features extracted from gyroscope finger tapping data and discerned patients with PD and atypical parkinsonisms and healthy controls in a multiclass setting with overall accuracy of 85.18% in a leave-one-out cross validation paradigm, confirming Hypothesis 3. Given that the input to the kNN model comprised a set of features selected algorithmically through a semi-greedy approach (two features extracted from the gyroscope mounted on the thumb, and four features from the gyroscope on the index finger) and giving some insight into the inner workings of the decision model, we also confirm Hypothesis 4.

Deep learning approaches, although state-of-the-art in the field of artificial intelligence, delivered subpar performance in classifying the observed neurological disorders, most likely due to insufficient amounts of data, even when supplemented with synthetically generated samples, and could not be used to support Hypothesis 3.

The described methodology using gyroscope-based kinematic analysis could be easily applied in laboratories world-wide for standardization of data collection protocols, opening the door to creating a large multi-center finger tapping kinematics database, which could be used to further refine the predictive power of artificial intelligence algorithms, and gaining new insights into the nature of motion disturbances, thus providing medical doctors with a quick clinical aid in differential diagnostics of the selected neurodegenerative disorders. The best use of the system would be to recruit as many early-stage participants as possible, as this is the time when clinical diagnosis is the most elusive in a more complex, longitudinal study, where *de novo* patients would be tested and recorded using the proposed system but would then go on to be followed for a number of years to make sure the correct diagnosis is reached with high certainty, damping the effect of possible noise in the labels (non-autopsy-confirmed diagnoses). Future work should also include turning the presented analyses into a user-friendly application that could be adopted in the clinical settings, through working closely with experts of clinical neurology.

As a support test for the posed hypotheses, this work also tackled PD diagnostics through the use of artificial intelligence in kinematic analysis of gait, using a sensorized electronic walkway, which provides a plethora of spatio-temporal features of gait, however without clear standardization on feature selection for detection of PD pathology. *De novo* PD patients and a healthy control group were recruited for this study and tested in a series of dual-task tests, where the interference task was of motor or mental type. Analyses relying on machine learning have been done on this set of gait data to test whether this modality can be used to aid in discerning PD patients from healthy controls (Hypothesis 2) and if a set of relevant parameters can be obtained programmatically that would be the most informative in aiding PD diagnostics (Hypothesis 4). This group of tests did not involve patients with atypical parkinsonisms, so the hypotheses that pertain to differentiation of PD from atypical parkinsonisms were not tested using gait kinematics, but only through finger motion.

With the help of clustering performed on GAITRite-provided features with inter-correlation as similarity metrics, and random forest feature importance ranking for further dimensionality reduction, we were able to programmatically select a subset of parameters that best help in PD diagnostics, and thus confirm Hypothesis 4, in the scope of differentiating PD from healthy controls. This subset of features included stride length and stride length CV, stride time and stride time CV, swing time and swing time CV, step time asymmetry, and heel-to-heel base support CV. The feature ranking algorithm showed the single most important variable for detecting *de novo* PD from gait was stride length variation during the mental task, emphasizing gait variability, particularly in the mental task, suggesting impairment of gait automation as an early sign of PD.

When the selected parameters were used as input to an SVM classifier, the achieved accuracy was 85%, confirming Hypothesis 2, that is, showing that AI and kinematic gait analysis can be used to help discern *de novo* PD patients from healthy controls on an individual level. High predictive power of the performance on gait in dual tasks was seen in the classifier results, where the mental task alone provided the highest results of all three (82% accuracy), followed closely by the motor task (81%) and the lowest scores were obtained for the base task (76%), although the combination of all three tasks outperformed each task taken individually (85%).

Future work should include recruiting patients with atypical parkinsonisms and using the described analysis of gait to test Hypothesis 1 and Hypothesis 4, i.e. whether statistical analyses of gait and artificial intelligence can help guide differential diagnostics between PD and atypical parkinsonisms on group and patient level.

This work shows the ability of artificial intelligence algorithms combined with motion sensors to aid in differential diagnostics of Parkinson's disease when compared to healthy individuals, as well as with

selected atypical parkinsonisms, contributing to the diagnostic arsenal used in clinical practice. Apart from giving blind diagnostic suggestions, artificial intelligence can be used to select a subset of features that bear the most predictive power in discerning PD from other pathologies, potentially adding weight to previous findings, or even pointing to new parameters of interest that have so far eluded the scientific efforts.

7. APPENDIX

7.1. Algorithm 1: Semi-greedy feature selection algorithm

Selected_set = { }

best_performance = 0

while *running time < Y* **do**

    //expansion phase

    Current_set = Selected_set

    for *i = 1: N* **do**

        //Change M features in Selected_set

        for *j = 1: M* **do**

            Randomly pick action from set {add, remove, swap}

            if *'add'* **do**

                Randomly select without replacement a feature from Remaining_features and add to Current_set

            else if *'remove'* **do**

                randomly remove a feature from Current_set

            else *//'substitute'*

                Randomly replace a feature from Current_set with a feature from Remaining_features

            end

        end

        test and save performance of Current_set

    end

    Expanded_set = union(K best performing sets)

    //reduction phase

    Reduced_set = { }

    for *i = 1: Nmax* **do**

        Feature_set_list = select combinations of

exactly i features from Expanded_set

for *Feature_set in Feature_set_list* **do**

test performance of Feature_set

if *performance > best_performance* **do**

Reduced_set = Feature_set

best_performance = performance

end

end

end

Selected_set = Reduced_set

end

7.2.   Code structure: multiclass classification

The code for feature extraction and classification of tapping data based on traditional machine learning approaches was written in Python 3.7.7. in an object-oriented paradigm, and relies on several major parts:

- Signal
- Artefact
- ArtefactExtractor
- ArtefactNormalizer
- ArtefactSelectorGenerator
- ArtefactFilter
- ArtefactTestGenerator
- ArtefactEvaluatorGenerator
- ArtefactEvaluator
- Parameters
- Utility functions

Signal is the class containing a recording and its metadata, including clinical diagnosis, date and time of recording, participant code, time of the first and last tap in the signal and other. Its methods allow for signal plotting, transformation into spherical coordinates, copying, and outputting a summary info.
All signals are loaded into their respective Signal objects. Then the ArtefactExtractor module is called to extract features from each measurement. It first applies a transformation function to the signal, and then extracts the features, as described in chapter 3.nestonesto.
ArtifactNormalizer is called to normalize each feature to the range [0,1].
Artefacts are optionally filtered by ArtefactFilter module, based on a manually chosen subset of features, and ArtefactSelectorGenerator handles the feature selection process.
ArtefactTestGenerator yields folds for leave-one-out cross validation.
ArtefactEvaluatorGenerator chooses the machine learning algorithm (or a number of algorithms) to fit and test, and provides a means to combine evaluations from all crossvalidation folds.
ArtefactEvaluator is called to evaluate the extracted features on all folds and give final results.

List of figures

List of tables

Used abbreviations

ACE-R - Addenbrooke's cognitive examination revised
Adam - Adaptive moment estimation optimizer
ANN - Artificial neural network
ANOVA - Analysis of variance
AS - Apathy scale
AU - Action unit
BDI - Beck Depression Inventory
BERT - Bidirectional Encoder Representations from Transformers
CNN - Convolutional neural network
CT - Computed tomography
CV - Coefficient of variation
DBS - Deep brain stimulation
DNN - Dynamic neural network
DT - Decision tree
DTI - Diffusion tensor imaging
DWT - Discrete wavelet transformation
EEG - Electro-encephalogram
EML - Extreme machine learning
ENS - Ensemble of classifiers
EVOL - Evolutionary algorithm
fMRI - functional magnetic resonance imaging
FoG - Freezing of gait
GABA - Gamma-aminobyturic acid
GAN - Generative adversarial network
GUI - Graphical user interface
H&Y - Hoehn and Yahr scale
HAM-A - Hamilton anxiety rating scale
HAM-D - Hamilton depression rating scale
HC - Healthy controls
H-H - Heel-to-heel

HMM - Hidden Markov model
IMU - inertial motion unit
kNN - k nearest neighbours
LDA - Linear Discriminant analysis
LDM - large margin distibution machine
LR - Logistic regression
MDS - Movement disorder society
MEMS - Micro electromechanical system
ML - Machine learning
MRI - Magnetic resonance imaging
MSA - Multiple System Atrophy
MSA-C - Multiple System Atrophy with predominant cerebellar ataxia
MSA-P - Multiplse System Atrophy with predominant parkinsonism
NB - Naive Bayes
NN - Neural network
NPV - negative predictive value
PC - Personal computer
PCA - Principal component analysis
PD - Parkinson's Disease
PDSBB - Parkinson's Disease Society Brain Bank
PNN - Probabilistic neural network
PPMI - parkinson's progression markers initiative
PPV - Positive predictive value
PSP - Progressive Supranuclear Palsy
RBM - Restricted Boltzman machine
ReLU - rectified linear unit
RF - Random forest
RMS - Root mean square
ROI - Region of interest
Se - Sensitivity
SNc - Substantia nigra pars compacta
SNr - Substantia nigra pars reticulata
Sp - Specificity
SPECT - Single photon emission tomography
STD - Standard deviation
STN - Subthalamic nucleus
SVM - Support vector machine
SVM RBF - Support vector machine with Radial basis kernel
SWI - Susceptibility weighted imaging
TCS - transcranial sonography
TREE - Tree based algorithms
UPDR - Unified Parkinson's Disease Rating Scale
WT - Wavelet transformation

References

[1]   J. Jankovic, "Parkinson's disease: clinical features and diagnosis," *J Neurol Neurosurg Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008, doi: 10.1136/JNNP.2007.131045.

[2]   ICD-11, "Parkinsonism," 2022. https://icd.who.int/browse11/l-m/en#/http%3a%2f%2fid.who.int%2ficd%2fentity%2f2024168133 (accessed Feb. 17, 2022).

[3]   C. G. Goetz, "The History of Parkinson's Disease: Early Clinical Descriptions and Neurological Therapies," *Cold Spring Harb Perspect Med*, vol. 1, no. 1, Sep. 2011, doi: 10.1101/CSHPERSPECT.A008862.

[4]   T. Pringsheim, N. Jette, A. Frolkis, and T. D. L. Steeves, "17. The prevalence of Parkinson's disease: A systematic review and meta-analysis," *Movement Disorders*, 2014, doi: 10.1002/mds.25945.

[5]   A. Rossi, K. Berger, H. Chen, D. Leslie, R. B. Mailman, and X. Huang, "Projection of the prevalence of Parkinson's disease in the coming decades: Revisited," *Movement Disorders*, 2018, doi: 10.1002/mds.27063.

[6]   J. Ebina, S. Ebihara, and O. Kano, "Similarities, differences and overlaps between frailty and Parkinson's disease," *Geriatr Gerontol Int*, Mar. 2022, doi: 10.1111/GGI.14362.

[7]   E. Tolosa, A. Garrido, S. W. Scholz, and W. Poewe, "Challenges in the diagnosis of Parkinson's disease," *Lancet Neurol*, vol. 20, no. 5, pp. 385–397, May 2021, doi: 10.1016/S1474-4422(21)00030-2.

[8]   C. H. Adler *et al.*, "Low clinical diagnostic accuracy of early vs advanced Parkinson disease: Clinicopathologic study," *Neurology*, 2014, doi: 10.1212/WNL.0000000000000641.

[9]   G. Rizzo, M. Copetti, S. Arcuti, D. Martino, A. Fontana, and G. Logroscino, "Accuracy of clinical diagnosis of Parkinson disease," *Neurology*, vol. 86, no. 6, pp. 566–576, Feb. 2016, doi: 10.1212/WNL.0000000000002350.

[10]   S. Grillner, B. Robertson, and M. Stephenson-Jones, "The evolutionary origin of the vertebrate basal ganglia and its role in action selection," *J Physiol*, vol. 591, no. 22, pp. 5425–5431, Nov. 2013, doi: 10.1113/JPHYSIOL.2012.246660.

[11]   T. Wichmann, H. Bergman, and M. R. DeLong, "Basal ganglia, movement disorders and deep brain stimulation: advances made through non-human primate research," *Journal of Neural Transmission 2017 125:3*, vol. 125, no. 3, pp. 419–430, Jun. 2017, doi: 10.1007/S00702-017-1736-5.

[12]   K. Simonyan, "Recent advances in understanding the role of the basal ganglia," *F1000Res*, vol. 8, 2019, doi: 10.12688/F1000RESEARCH.16524.1.

[13]   A. M. Graybiel, "The basal ganglia," *Current Biology*, vol. 10, no. 14, pp. R509–R511, Jul. 2000, doi: 10.1016/S0960-9822(00)00593-5.

[14]   E. M. Adam, E. N. Brown, N. Kopell, and M. M. McCarthy, "Deep brain stimulation in the subthalamic nucleus for Parkinson's disease can restore dynamics of striatal networks," *Proc Natl Acad Sci U S A*, vol. 119, no. 19, p. e2120808119, May 2022, doi: 10.1073/PNAS.2120808119/SUPPL_FILE/PNAS.2120808119.SAPP.PDF.

[15]   A. Bartels and S. Zeki, "The neural basis of romantic love," *NeuroReport*, vol. 11, no. 17, pp. 3829–3834, Nov. 2000.

[16] M. D. Sacchet, M. C. Camacho, E. E. Livermore, E. A. C. Thomas, and I. H. Gotlib, "Accelerated aging of the putamen in patients with major depressive disorder," *Journal of Psychiatry and Neuroscience*, vol. 42, no. 3, pp. 164–171, May 2017, doi: 10.1503/JPN.160010.

[17] D. W. Dickson *et al.*, "Neuropathological assessment of Parkinson's disease: refining the diagnostic criteria," *Lancet Neurol*, vol. 8, no. 12, pp. 1150–1157, Dec. 2009, doi: 10.1016/S1474-4422(09)70238-8.

[18] H. Braak, R. A. I. de Vos, J. Bohl, and K. del Tredici, "Gastric α-synuclein immunoreactive inclusions in Meissner's and Auerbach's plexuses in cases staged for Parkinson's disease-related brain pathology," *Neurosci Lett*, vol. 396, no. 1, pp. 67–72, Mar. 2006, doi: 10.1016/J.NEULET.2005.11.012.

[19] L. Stefanis, "α-Synuclein in Parkinson's Disease," *Cold Spring Harb Perspect Med*, vol. 2, no. 2, 2012, doi: 10.1101/CSHPERSPECT.A009399.

[20] M. E. Morris, R. Iansek, T. A. Matyas, and J. J. Summers, "The pathogenesis of gait hypokinesia in Parkinson's disease," *Brain*, vol. 117 (Pt 5), no. 5, pp. 1169–1181, Oct. 1994, doi: 10.1093/BRAIN/117.5.1169.

[21] G. Yogev-Seligmann, J. M. Hausdorff, and N. Giladi, "The role of executive function and attention in gait," *Mov Disord*, vol. 23, no. 3, pp. 329–342, Feb. 2008, doi: 10.1002/MDS.21720.

[22] B. R. Bloem, Y. A. M. Grimbergen, J. G. van Dijk, and M. Munneke, "The 'posture second' strategy: A review of wrong priorities in Parkinson's disease," *J Neurol Sci*, vol. 248, no. 1–2, pp. 196–204, Oct. 2006, doi: 10.1016/j.jns.2006.05.010.

[23] T. Wu and M. Hallett, "A functional MRI study of automatic movements in patients with Parkinson's disease," *Brain*, vol. 128, no. 10, pp. 2250–2259, Oct. 2005, doi: 10.1093/BRAIN/AWH569.

[24] R. Baltadjieva, N. Giladi, L. Gruendlinger, C. Peretz, and J. M. Hausdorff, "Marked alterations in the gait timing and rhythmicity of patients with *de novo* Parkinson's disease," *European Journal of Neuroscience*, vol. 24, no. 6, pp. 1815–1820, Sep. 2006, doi: 10.1111/J.1460-9568.2006.05033.X.

[25] C. Gao, J. Liu, Y. Tan, and S. Chen, "Freezing of gait in Parkinson's disease: pathophysiology, risk factors and treatments," *Translational Neurodegeneration 2020 9:1*, vol. 9, no. 1, pp. 1–22, Apr. 2020, doi: 10.1186/S40035-020-00191-5.

[26] T. Gazibara *et al.*, "Fall frequency and risk factors in patients with Parkinson's disease in Belgrade, Serbia: A cross-sectional study," *Geriatr Gerontol Int*, 2015, doi: 10.1111/ggi.12300.

[27] P. Wu *et al.*, "Objectifying facial expressivity assessment of parkinson's patients: Preliminary study," *Comput Math Methods Med*, vol. 2014, 2014, doi: 10.1155/2014/427826.

[28] M. T. M. Prenger and P. A. Macdonald, "Problems with Facial Mimicry Might Contribute to Emotion Recognition Impairment in Parkinson's Disease," *Parkinsons Dis*, vol. 2018, 2018, doi: 10.1155/2018/5741941.

[29] S. Argaud, M. Vérin, P. Sauleau, and D. Grandjean, "Facial emotion recognition in Parkinson's disease: A review and new hypotheses," *Movement Disorders*, vol. 33, no. 4, pp. 554–567, Apr. 2018, doi: 10.1002/MDS.27305.

[30] P. Jacob, "A Philosopher's Reflections on the Discovery of Mirror Neurons," *Top Cogn Sci*, vol. 1, no. 3, pp. 570–595, Jul. 2009, doi: 10.1111/J.1756-8765.2009.01040.X.

[31] P. Molenberghs, R. Cunnington, and J. B. Mattingley, "Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies," *Neurosci Biobehav Rev*, vol. 36, no. 1, pp. 341–349, Jan. 2012, doi: 10.1016/J.NEUBIOREV.2011.07.004.

[32] M. Thomas, A. Lenka, and P. Kumar Pal, "Handwriting Analysis in Parkinson's Disease: Current Status and Future Directions," *Mov Disord Clin Pract*, vol. 4, no. 6, pp. 806–818, Nov. 2017, doi: 10.1002/MDC3.12552.

[33] 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, doi: 10.1002/MDS.22340.

[34] D. W. Dickson, "Parkinson's Disease and Parkinsonism: Neuropathology," *Cold Spring Harb Perspect Med*, vol. 2, no. 8, p. a009258, Aug. 2012, doi: 10.1101/CSHPERSPECT.A009258.

[35] J. C. Steele, J. C. Richardson, and J. Olszewski, "Progressive Supranuclear Palsy: A Heterogeneous Degeneration Involving the Brain Stem, Basal Ganglia and Cerebellum With Vertical Gaze and Pseudobulbar Palsy, Nuchal Dystonia and Dementia," *Arch Neurol*, vol. 10, no. 4, pp. 333–359, Apr. 1964, doi: 10.1001/ARCHNEUR.1964.00460160003001.

[36] L. I. Golbe, "Progressive supranuclear palsy," *Semin Neurol*, vol. 34, no. 2, pp. 151–159, Jun. 2014, doi: 10.1055/S-0034-1381736/ID/JR00925-38.

[37] H. C. Shi *et al.*, "Gray matter atrophy in progressive supranuclear palsy: meta-analysis of voxel-based morphometry studies," *Neurological Sciences*, vol. 34, no. 7, pp. 1049–1055, 2013.

[38] G. U. Höglinger *et al.*, "Clinical diagnosis of progressive supranuclear palsy: The movement disorder society criteria," *Movement Disorders*, 2017, doi: 10.1002/mds.26987.

[39] A. E. Lang, "Treatment of progressive supranuclear palsy and corticobasal degeneration," *Mov Disord*, vol. 20, no. S12, pp. S83--S91, 2005.

[40] R. Constantinescu, I. Richard, and R. Kurlan, "Levodopa responsiveness in disorders with parkinsonism: A review of the literature," *Movement Disorders*. 2007. doi: 10.1002/mds.21578.

[41] S. Koga *et al.*, "When DLB, PD, and PSP masquerade as MSA: an autopsy study of 134 patients.," *Neurology*, 2015, doi: 10.1212/WNL.0000000000001807.

[42] W. G. Meissner *et al.*, "Multiple System Atrophy: Recent Developments and Future Perspectives," *Movement Disorders*, vol. 34, no. 11, pp. 1629–1642, Nov. 2019, doi: 10.1002/MDS.27894.

[43] A. J. Hughes, S. E. Daniel, L. Kilford, and A. J. Lees, "Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases," *J Neurol Neurosurg Psychiatry*, vol. 55, no. 3, pp. 181–184, 1992, doi: 10.1136/JNNP.55.3.181.

[44] 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, 2008, doi: 10.1002/mds.22340.

[45] "MDS-UPDRS." https://www.movementdisorders.org/MDS/MDS-Rating-Scales/MDS-Unified-Parkinsons-Disease-Rating-Scale-MDS-UPDRS.htm (accessed Jul. 15, 2022).

[46]   C. C. Lainscsek *et al.*, "Finger tapping movements of Parkinson's disease patients automatically rated using nonlinear delay differential equations," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 1, p. 013119, Feb. 2012, doi: 10.1063/1.3683444.

[47]   R. Bhidayasiri and D. Tarsy, "Parkinson's Disease: Hoehn and Yahr Scale," *Current Clinical Neurology*, vol. 36, pp. 4–5, 2012, doi: 10.1007/978-1-60327-426-5_2/COVER/.

[48]   C. G. Goetz *et al.*, "Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations," *Mov Disord*, vol. 19, no. 9, pp. 1020–1028, Sep. 2004, doi: 10.1002/MDS.20213.

[49]   C. P. Weingarten, M. H. Sundman, P. Hickey, and N. kuei Chen, "Neuroimaging of Parkinson's disease: Expanding views," *Neuroscience and Biobehavioral Reviews*. 2015. doi: 10.1016/j.neubiorev.2015.09.007.

[50]   C. Salvatore *et al.*, "Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy," *J Neurosci Methods*, 2014, doi: 10.1016/j.jneumeth.2013.11.016.

[51]   N. K. Focke *et al.*, "Individual voxel-based subtype prediction can differentiate progressive supranuclear palsy from idiopathic Parkinson syndrome and healthy controls," *Hum Brain Mapp*, 2011, doi: 10.1002/hbm.21161.

[52]   S. Duchesne, Y. Rolland, and M. Vérin, "Automated Computer Differential Classification in Parkinsonian Syndromes via Pattern Analysis on MRI," *Acad Radiol*, 2009, doi: 10.1016/j.acra.2008.05.024.

[53]   C. Scherfler *et al.*, "Diagnostic potential of automated subcortical volume segmentation in atypical parkinsonism," *Neurology*, 2016, doi: 10.1212/WNL.0000000000002518.

[54]   A. F. Marquand *et al.*, "Automated, High Accuracy Classification of Parkinsonian Disorders: A Pattern Recognition Approach," *PLoS One*, 2013, doi: 10.1371/journal.pone.0069237.

[55]   S. Soltaninejad, I. Cheng, and A. Basu, "Towards the identification of Parkinson's Disease using only T1 MR Images," *arXiv preprint arXiv:1806.07489*, 2018.

[56]   B. Rana *et al.*, "Regions-of-interest based automated diagnosis of Parkinson's disease using T1-weighted MRI," *Expert Syst Appl*, 2015, doi: 10.1016/j.eswa.2015.01.062.

[57]   B. Peng *et al.*, "A multilevel-ROI-features-based machine learning method for detection of morphometric biomarkers in Parkinson's disease," *Neurosci Lett*, 2017, doi: 10.1016/j.neulet.2017.04.034.

[58]   G. Singh, M. Vadera, L. Samavedham, and E. C. H. Lim, "Machine Learning-Based Framework for Multi-Class Diagnosis of Neurodegenerative Diseases: A Study on Parkinson's Disease," *IFAC-PapersOnLine*, 2016, doi: 10.1016/j.ifacol.2016.07.331.

[59]   S. Haller, S. Badoud, D. Nguyen, V. Garibotto, K. O. Lovblad, and P. R. Burkhard, "Individual detection of patients with Parkinson Disease using support vector machine analysis of diffusion tensor imaging data: Initial results," *American Journal of Neuroradiology*, 2012, doi: 10.3174/ajnr.A3126.

[60]   P. J. Planetta *et al.*, "Free-water imaging in Parkinson's disease and atypical parkinsonism," *Brain*, 2016, doi: 10.1093/brain/awv361.

[61] S. Haller *et al.*, "Differentiation between Parkinson disease and other forms of Parkinsonism using support vector machine analysis of susceptibility-weighted imaging (SWI): Initial results," *Eur Radiol*, 2013, doi: 10.1007/s00330-012-2579-y.

[62] A. Cherubini *et al.*, "Magnetic resonance support vector machine discriminates between Parkinson disease and progressive supranuclear palsy," *Movement Disorders*, vol. 29, no. 2, pp. 266–269, 2014.

[63] P. Péran *et al.*, "MRI supervised and unsupervised classification of Parkinson's disease and multiple system atrophy," *Movement Disorders*, 2018, doi: 10.1002/mds.27307.

[64] H. J. Huppertz *et al.*, "Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification," *Movement Disorders*, 2016, doi: 10.1002/mds.26715.

[65] R. Morisi *et al.*, "Multi-class parkinsonian disorders classification with quantitative MR markers and graph-based features using support vector machines," *Parkinsonism Relat Disord*, 2018, doi: 10.1016/j.parkreldis.2017.11.343.

[66] H. Choi, S. Ha, H. J. Im, S. H. Paek, and D. S. Lee, "Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging," *Neuroimage Clin*, 2017, doi: 10.1016/j.nicl.2017.09.010.

[67] E. Adeli, G. Wu, B. Saghafi, L. An, F. Shi, and D. Shen, "Kernel-based Joint Feature Selection and Max-Margin Classification for Early Diagnosis of Parkinson's Disease," *Sci Rep*, 2017, doi: 10.1038/srep41069.

[68] M. Rumman, A. N. Tasneem, S. Farzana, and others, "Early detection of Parkinson's disease using image processing and artificial neural network," BRAC University, 2018.

[69] D. H. Kim, H. Wit, and M. Thurston, "Artificial intelligence in the diagnosis of Parkinson's disease from ioflupane-123 single-photon emission computed tomography dopamine transporter scans using transfer learning," *Nucl Med Commun*, vol. 39, no. 10, pp. 887–893, 2018.

[70] N. Amoroso, M. la Rocca, A. Monaco, R. Bellotti, and S. Tangaro, "Complex networks reveal early MRI markers of Parkinson's disease," *Med Image Anal*, 2018, doi: 10.1016/j.media.2018.05.004.

[71] S. Esmaeilzadeh, Y. Yang, and E. Adeli, "End-to-End Parkinson Disease Diagnosis using Brain MR-Images by 3D-CNN," *arXiv preprint arXiv:1806.05233*, 2018.

[72] B. Gong *et al.*, "Neuroimaging-based diagnosis of Parkinson's disease with deep neural mapping large margin distribution machine," *Neurocomputing*, 2018. doi: 10.1016/j.neucom.2018.09.025.

[73] T. Zhang and Z. H. Zhou, "Large margin distribution machine," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 313–322, 2014, doi: 10.1145/2623330.2623710.

[74] K. Marek *et al.*, "The Parkinson Progression Marker Initiative (PPMI)," *Prog Neurobiol*, vol. 95, no. 4, pp. 629–635, Dec. 2011, doi: 10.1016/J.PNEUROBIO.2011.09.005.

[75] E. Rovini, C. Maremmani, and F. Cavallo, "How wearable sensors can support parkinson's disease diagnosis and treatment: A systematic review," *Frontiers in Neuroscience*. 2017. doi: 10.3389/fnins.2017.00555.

[76] C. Ahlrichs and M. Lawo, "Parkinson's Disease Motor Symptoms in Machine Learning: A Review," *Health Informatics - An International Journal*, vol. 2, no. 4, pp. 1–18, Dec. 2013, doi: 10.5121/hiij.2013.2401.

[77] L. Giancardo *et al.*, "Computer keyboard interaction as an indicator of early Parkinson's disease," *Sci Rep*, 2016, doi: 10.1038/srep34468.

[78] W. R. Adams, "High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing," *PLoS One*, 2017, doi: 10.1371/journal.pone.0188226.

[79] T. Arroyo-Gallego *et al.*, "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," *J Med Internet Res*, 2018, doi: 10.2196/jmir.9462.

[80] M. D. Djurić-Jovičić *et al.*, "Implementation of continuous wavelet transformation in repetitive finger tapping analysis for patients with PD," in *2014 22nd Telecommunications Forum, TELFOR 2014 - Proceedings of Papers*, 2014. doi: 10.1109/TELFOR.2014.7034466.

[81] L. Fraiwan, R. Khnouf, and A. R. Mashagbeh, "Parkinsons disease hand tremor detection system for mobile application," *J Med Eng Technol*, 2016, doi: 10.3109/03091902.2016.1148792.

[82] D. Iakovakis, S. Hadjidimitriou, V. Charisis, S. Bostantzopoulou, Z. Katsarou, and L. J. Hadjileontiadis, "Touchscreen typing-pattern analysis for detecting fine motor skills decline in early-stage Parkinson's disease," *Sci Rep*, 2018, doi: 10.1038/s41598-018-25999-0.

[83] N. Kostikis, D. Hristu-Varsakelis, M. Arnaoutoglou, and C. Kotsavasiloglou, "A smartphone-based tool for assessing parkinsonian hand tremor," *IEEE J Biomed Health Inform*, 2015, doi: 10.1109/JBHI.2015.2471093.

[84] M. A. Lones *et al.*, "Evolving classifiers to recognize the movement characteristics of parkinson's disease patients," *IEEE Transactions on Evolutionary Computation*, 2014, doi: 10.1109/TEVC.2013.2281532.

[85] T. Khan, D. Nyholm, J. Westin, and M. Dougherty, "A computer vision framework for finger-tapping evaluation in Parkinson's disease," *Artif Intell Med*, 2014, doi: 10.1016/j.artmed.2013.11.004.

[86] C. G. Goetz *et al.*, "Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): Process, format, and clinimetric testing plan," *Movement Disorders*, vol. 22, no. 1, pp. 41–47, Jan. 2007, doi: 10.1002/MDS.21198.

[87] Y. Nancy Jane, H. Khanna Nehemiah, and K. Arputharaj, "A Q-backpropagated time delay neural network for diagnosing severity of gait disturbances in Parkinson's disease," *J Biomed Inform*, 2016, doi: 10.1016/j.jbi.2016.01.014.

[88] M. R. Daliri, "Chi-square distance kernel of the gaits for the diagnosis of Parkinson's disease," *Biomed Signal Process Control*, 2013, doi: 10.1016/j.bspc.2012.04.007.

[89] W. Zeng, F. Liu, Q. Wang, Y. Wang, L. Ma, and Y. Zhang, "Parkinson's disease classification using gait analysis via deterministic learning," *Neurosci Lett*, 2016, doi: 10.1016/j.neulet.2016.09.043.

[90] I. Arcolin, S. Corna, M. Giardini, A. Giordano, A. Nardone, and M. Godi, "Proposal of a new conceptual gait model for patients with Parkinson's disease based on factor analysis," *Biomed Eng Online*, vol. 18, no. 1, pp. 1–18, Jun. 2019, doi: 10.1186/S12938-019-0689-3/TABLES/3.

[91]   C. Tucker *et al.*, "A data mining methodology for predicting early stage Parkinson's disease using non-invasive, high-dimensional gait sensor data," *IIE Trans Healthc Syst Eng*, 2015, doi: 10.1080/19488300.2015.1095256.

[92]   D. J. Cook, M. Schmitter-Edgecombe, and P. Dawadi, "Analyzing activity behavior and movement in a naturalistic environment using smart home techniques," *IEEE J Biomed Health Inform*, 2015, doi: 10.1109/JBHI.2015.2461659.

[93]   A. Procházka, O. Vyšata, M. Vališ, O˛upa, M. Schätz, and V. Mařík, "Bayesian classification and analysis of gait disorders using image and depth sensors of Microsoft Kinect," *Digital Signal Processing: A Review Journal*, 2015, doi: 10.1016/j.dsp.2015.05.011.

[94]   F. Wahid, R. K. Begg, C. J. Hass, S. Halgamuge, and D. C. Ackland, "Classification of Parkinson's disease gait using spatial-temporal gait features," *IEEE J Biomed Health Inform*, vol. 19, no. 6, pp. 1794–1802, Nov. 2015, doi: 10.1109/JBHI.2015.2450232.

[95]   S. Arora, V. Venkataraman, S. Donohue, K. M. Biglan, E. R. Dorsey, and M. A. Little, "High accuracy discrimination of Parkinson's disease participants from healthy controls using smartphones," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2014. doi: 10.1109/ICASSP.2014.6854280.

[96]   Q. W. Oung, M. Hariharan, H. L. Lee, S. N. Basah, M. Sarillee, and C. H. Lee, "Wearable multimodal sensors for evaluation of patients with Parkinson disease," in *Proceedings - 5th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2015*, 2016. doi: 10.1109/ICCSCE.2015.7482196.

[97]   J. Barth *et al.*, "Biometric and mobile gait analysis for early diagnosis and therapy monitoring in Parkinson's disease," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2011. doi: 10.1109/IEMBS.2011.6090226.

[98]   J. Klucken *et al.*, "Unbiased and Mobile Gait Analysis Detects Motor Impairment in Parkinson's Disease," *PLoS One*, 2013, doi: 10.1371/journal.pone.0056956.

[99]   Q. W. Oung, H. Muthusamy, S. N. Basah, H. Lee, and V. Vijean, "Empirical Wavelet Transform Based Features for Classification of Parkinson's Disease Severity," *J Med Syst*, vol. 42, no. 2, p. 29, 2018.

[100]   C. Caramia *et al.*, "IMU-Based Classification of Parkinson's Disease From Gait: A Sensitivity Analysis on Sensor Location and Feature Selection," *IEEE J Biomed Health Inform*, vol. 22, no. 6, pp. 1765–1774, 2018.

[101]   F. Cuzzolin *et al.*, "Metric learning for Parkinsonian identification from IMU gait measurements," *Gait Posture*, 2017, doi: 10.1016/j.gaitpost.2017.02.012.

[102]   VectorNav, "Inertial sensors," Feb. 17, 2022. https://www.vectornav.com/resources/inertial-navigation-primer/theory-of-operation/theory-inertial (accessed Feb. 17, 2022).

[103]   D. K. Shaeffer, "MEMS inertial sensors: A tutorial overview," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 100–109, 2013, doi: 10.1109/MCOM.2013.6495768.

[104]   B. P. Printy *et al.*, "Smartphone application for classification of motor impairment severity in Parkinson's disease," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 2014. doi: 10.1109/EMBC.2014.6944176.

[105] M. Belić, V. Bobić, M. Badža, N. Šolaja, M. Đurić-Jovičić, and V. S. Kostić, "Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—A review," *Clin Neurol Neurosurg*, 2019, doi: 10.1016/j.clineuro.2019.105442.

[106] O. Bazgir, S. A. H. Habibi, L. Palma, P. Pierleoni, and S. Nafees, "A Classification System for Assessment and Home Monitoring of Tremor in Patients with Parkinson's Disease," *J Med Signals Sens*, vol. 8, no. 2, p. 65, 2018.

[107] C. Stamate *et al.*, "Deep learning Parkinson's from smartphone data," in *2017 IEEE International Conference on Pervasive Computing and Communications, PerCom 2017*, 2017. doi: 10.1109/PERCOM.2017.7917848.

[108] N. Y. Hammerla, J. M. Fisher, P. Andras, L. Rochester, R. Walker, and T. Plötz, "PD Disease State Assessment in Naturalistic Environments using Deep Learning," *Aaai*, 2015, doi: 10.1007/s00128-007-9129-3.

[109] J. M. Fisher, N. Y. Hammerla, T. Ploetz, P. Andras, L. Rochester, and R. W. Walker, "Unsupervised home monitoring of Parkinson's disease motor symptoms using body-worn accelerometers," *Parkinsonism Relat Disord*, 2016, doi: 10.1016/j.parkreldis.2016.09.009.

[110] M. Memedi *et al.*, "Automatic Spiral Analysis for Objective Assessment of Motor Symptoms in Parkinson's Disease," *Sensors (Basel)*, 2015, doi: 10.3390/s150923727.

[111] B. M. Eskofier *et al.*, "Recent machine learning advancements in sensor-based mobility analysis: Deep learning for Parkinson's disease assessment," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2016-October, pp. 655–658, Oct. 2016, doi: 10.1109/EMBC.2016.7590787.

[112] B. Sijobert, M. Benoussaad, J. Denys, R. Pissard-Gibollet, C. Geny, and C. A. Coste, "Implementation and Validation of a Stride Length Estimation Algorithm, Using a Single Basic Inertial Sensor on Healthy Subjects and Patients Suffering from Parkinson's Disease," *ElectronicHealthcare*, vol. 07, no. 06, pp. 704–714, 2015, doi: 10.4236/HEALTH.2015.76084.

[113] M. Djurić-Jovičić, M. Belić, I. Stanković, S. Radovanović, and V. S. Kostić, "Selection of gait parameters for differential diagnostics of patients with *de novo* Parkinson's disease," *Neurol Res*, 2017, doi: 10.1080/01616412.2017.1348690.

[114] K. J. Kubota, J. A. Chen, and M. A. Little, "Machine learning for large-scale wearable sensor data in Parkinson's disease: Concepts, promises, pitfalls, and futures," *Movement Disorders*, vol. 31, no. 9, pp. 1314–1326, Sep. 2016, doi: 10.1002/MDS.26693.

[115] C. Ahlrichs *et al.*, "Detecting freezing of gait with a tri-axial accelerometer in Parkinson's disease patients," *Med Biol Eng Comput*, 2016, doi: 10.1007/s11517-015-1395-3.

[116] D. Rodríguez-Martín *et al.*, "Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer," *PLoS One*, 2017, doi: 10.1371/journal.pone.0171764.

[117] N. K. Orphanidou, A. Hussain, R. Keight, P. Lishoa, J. Hind, and H. Al-Askar, "Predicting Freezing of Gait in Parkinsons Disease Patients Using Machine Learning," in *2018 IEEE Congress on Evolutionary Computation (CEC)*, 2018, pp. 1–8.

[118] E. E. Tripoliti *et al.*, "Automatic detection of freezing of gait events in patients with Parkinson's disease," *Comput Methods Programs Biomed*, 2013, doi: 10.1016/j.cmpb.2012.10.016.

[119] A. T. Tzallas *et al.*, "Perform: A system for monitoring, Assessment and management of patients with Parkinson's disease," *Sensors (Switzerland)*, 2014, doi: 10.3390/s141121329.

[120] S. Arora *et al.*, "Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study," *Parkinsonism Relat Disord*, vol. 21, no. 6, pp. 650–653, Jun. 2015, doi: 10.1016/J.PARKRELDIS.2015.02.026.

[121] V. Sharma *et al.*, "Spark: Personalized parkinson disease interventions through synergy between a smartphone and a smartwatch," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2014. doi: 10.1007/978-3-319-07635-5_11.

[122] S. Mazilu, A. Calatroni, E. Gazit, D. Roggen, J. M. Hausdorff, and G. Tröster, "Feature learning for detection and prediction of freezing of gait in Parkinson's disease," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2013. doi: 10.1007/978-3-642-39712-7_11.

[123] R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," *BioMedical Engineering Online*. 2013. doi: 10.1186/1475-925X-12-66.

[124] Y. Nancy Jane, H. Khanna Nehemiah, and K. Arputharaj, "A Q-backpropagated time delay neural network for diagnosing severity of gait disturbances in Parkinson's disease," *J Biomed Inform*, vol. 60, pp. 169–176, Apr. 2016, doi: 10.1016/J.JBI.2016.01.014.

[125] A. Rodríguez-Molinero *et al.*, "Validation of a Portable Device for Mapping Motor and Gait Disturbances in Parkinson's Disease," *JMIR Mhealth Uhealth 2015;3(1):e9 https://mhealth.jmir.org/2015/1/e9*, vol. 3, no. 1, p. e3321, Feb. 2015, doi: 10.2196/MHEALTH.3321.

[126] A. M. S. Muniz *et al.*, "Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect of subthalamic stimulation in Parkinson disease on ground reaction force during gait," *J Biomech*, vol. 43, no. 4, pp. 720–726, Mar. 2010, doi: 10.1016/J.JBIOMECH.2009.10.018.

[127] S. H. Roy *et al.*, "High-resolution tracking of motor disorders in Parkinson's disease during unconstrained activity," *Movement Disorders*, vol. 28, no. 8, pp. 1080–1087, Jul. 2013, doi: 10.1002/MDS.25391.

[128] S. Arora *et al.*, "Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study," *Parkinsonism Relat Disord*, 2015, doi: 10.1016/j.parkreldis.2015.02.026.

[129] S. Patel *et al.*, "Monitoring motor fluctuations in patients with parkinsons disease using wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 6, pp. 864–873, Nov. 2009, doi: 10.1109/TITB.2009.2033471.

[130] G. Rigas *et al.*, "Assessment of tremor activity in the parkinsons disease using a set of wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 478–487, 2012, doi: 10.1109/TITB.2011.2182616.

[131] M. Ashfak Habib, M. S. Mohktar, S. Bahyah Kamaruzzaman, K. Seang Lim, T. Maw Pin, and F. Ibrahim, "Smartphone-based solutions for fall detection and prevention: Challenges and open issues," *Sensors (Switzerland)*. 2014. doi: 10.3390/s140407181.

[132] C. Gao *et al.*, "Objective assessment of bradykinesia in Parkinson's disease using evolutionary algorithms: Clinical validation," *Transl Neurodegener*, 2018, doi: 10.1186/s40035-018-0124-x.

[133] R. Deb, S. An, G. Bhat, H. Shill, and U. Y. Ogras, "A Systematic Survey of Research Trends in Technology Usage for Parkinson's Disease," *Sensors (Basel)*, vol. 22, no. 15, Aug. 2022, doi: 10.3390/S22155491.

[134] A. M. A. Handojoseno *et al.*, "Prediction of freezing of gait using analysis of brain effective connectivity," *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, pp. 4119–4122, Nov. 2014, doi: 10.1109/EMBC.2014.6944530.

[135] A. M. Handojoseno *et al.*, "Prediction of Freezing of Gait in Patients with Parkinson's Disease Using EEG Signals.," *Stud Health Technol Inform*, vol. 246, pp. 124–131, 2018.

[136] N. Betrouni *et al.*, "Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: Preliminary results," *Movement Disorders*, 2018.

[137] A. Nogales, A. J. Garcia-Tejedor, A. M. Maitin, A. Perez-Morales, M. D. del Castillo, and J. P. Romero, "BERT learns from electroencephalograms about Parkinson´s Disease: Transformer-based models for aid diagnosis," *IEEE Access*, pp. 1–1, Aug. 2022, doi: 10.1109/ACCESS.2022.3201843.

[138] A. Joshi, L. Tickle-Degnen, S. Gunnery, T. Ellis, and M. Betke, "Predicting active facial expressivity in people with Parkinson's disease," *ACM International Conference Proceeding Series*, vol. 29-June-2016, Jun. 2016, doi: 10.1145/2910674.2910686.

[139] A. Bandini *et al.*, "Analysis of facial expressions in parkinson's disease through video-based automatic methods," *J Neurosci Methods*, vol. 281, pp. 7–20, Apr. 2017, doi: 10.1016/J.JNEUMETH.2017.02.006.

[140] A. Joshi, S. Ghosh, S. Gunnery, L. Tickle-Degnen, S. Sclaroff, and M. Betke, "Context-sensitive prediction of facial expressivity using multimodal hierarchical Bayesian neural networks," *Proceedings - 13th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2018*, pp. 278–285, Jun. 2018, doi: 10.1109/FG.2018.00048.

[141] N. Vinokurov, D. Arkadir, E. Linetsky, H. Bergman, and D. Weinshall, "Quantifying Hypomimia in Parkinson Patients Using a Depth Camera," *Communications in Computer and Information Science*, vol. 604, pp. 63–71, 2015, doi: 10.1007/978-3-319-32270-4_7.

[142] M. Djuric-Jovicic *et al.*, "Finger and foot tapping sensor system for objective motor assessment," *Vojnosanit Pregl*, 2018, doi: 10.2298/VSP150502323D.

[143] K. V. S. Djurić-Jovičić Milica, Petrović Igor, Ječmenica-Lukić Milica, Radovanović Saša, Dragašević-Mišković Nataša, Belić Minja, Miler-Jerković Vera, Popović Mirjana B., "Finger tapping analysis in patients with Parkinson's disease and atypical parkinsonism," *Journal of Clinical Neuroscience*, vol. 30, no. August 2016, pp. 49–55, 2016, doi: 10.1016/j.jocn.2015.10.053.

[144] H. Ling, L. A. Massey, A. J. Lees, P. Brown, and B. L. Day, "Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease," *Brain*, vol. 135, no. 4, pp. 1141–1153, Apr. 2012, doi: 10.1093/BRAIN/AWS038.

[145] S. Y. Kang *et al.*, "Characteristics of the sequence effect in Parkinson's disease," *Movement Disorders*, vol. 25, no. 13, pp. 2148–2155, Oct. 2010, doi: 10.1002/MDS.23251.

[146] M. Desmurget, S. T. Grafton, P. Vindras, H. Gréa, and R. S. Turner, "The basal ganglia network mediates the planning of movement amplitude," *European Journal of Neuroscience*, vol. 19, no. 10, pp. 2871–2880, May 2004, doi: 10.1111/J.0953-816X.2004.03395.X.

[147] M. M. Koop, B. C. Hill, and H. M. Bronte-Stewart, "Perceptual errors increase with movement duration and may contribute to hypokinesia in Parkinson's disease," *Neuroscience*, vol. 243, pp. 1–13, Jul. 2013, doi: 10.1016/J.NEUROSCIENCE.2013.03.026.

[148] E. Lee *et al.*, "Neural correlates of progressive reduction of bradykinesia in *de novo* Parkinson's disease," *Parkinsonism Relat Disord*, vol. 20, no. 12, pp. 1376–1381, Dec. 2014, doi: 10.1016/J.PARKRELDIS.2014.09.027.

[149] V. Pichot *et al.*, "Wavelet transform to quantify heart rate variability and to assess its instantaneous changes," *J Appl Physiol*, vol. 86, no. 3, pp. 1081–1091, 1999, doi: 10.1152/JAPPL.1999.86.3.1081/ASSET/IMAGES/LARGE/JAPP05304008X.JPEG.

[150] S. Thurner, M. C. Feurstein, and M. C. Teich, "Multiresolution Wavelet Analysis of Heartbeat Intervals Discriminates Healthy Patients from Those with Cardiac Pathology," *Phys Rev Lett*, vol. 80, no. 7, p. 1544, Feb. 1998, doi: 10.1103/PhysRevLett.80.1544.

[151] N. Wang, E. Ambikairajah, N. H. Lovell, and B. G. Celler, "Accelerometry based classification of walking patterns using time-frequency analysis," *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp. 4899–4902, 2007, doi: 10.1109/IEMBS.2007.4353438.

[152] M. N. Nyan, F. E. H. Tay, K. H. W. Seah, and Y. Y. Sitoh, "Classification of gait patterns in the time–frequency domain," *J Biomech*, vol. 39, no. 14, pp. 2647–2656, Jan. 2006, doi: 10.1016/J.JBIOMECH.2005.08.014.

[153] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Trans Biomed Eng*, vol. 56, no. 3, pp. 871–879, Mar. 2009, doi: 10.1109/TBME.2008.2006190.

[154] B. Ayrulu-Erdem and B. Barshan, "Leg Motion Classification with Artificial Neural Networks Using Wavelet-Based Features of Gyroscope Signals," *Sensors 2011, Vol. 11, Pages 1721-1743*, vol. 11, no. 2, pp. 1721–1743, Jan. 2011, doi: 10.3390/S110201721.

[155] J. Chakraborty and A. Nandy, "Discrete wavelet transform based data representation in deep neural network for gait abnormality detection," *Biomed Signal Process Control*, vol. 62, p. 102076, Sep. 2020, doi: 10.1016/J.BSPC.2020.102076.

[156] B. Ando *et al.*, "A Wavelet-Based Methodology for Features Extraction in Postural Instability Analysis," *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, vol. 2021-May, May 2021, doi: 10.1109/I2MTC50364.2021.9459816.

[157] P. Chaovalit, A. Gangopadhyay, G. Karabatis, and Z. Chen, "Discrete wavelet transform-based time series analysis and mining," *ACM Comput Surv*, vol. 43, no. 2, 2011, doi: 10.1145/1883612.1883613.

[158] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans Pattern Anal Mach Intell*, vol. 11, no. 7, 1989, doi: 10.1109/34.192463.

[159] "Daubechies 4 wavelet (db4) properties, filters and functions - Wavelet Properties Browser." http://wavelets.pybytes.com/wavelet/db4/ (accessed Aug. 30, 2022).

[160] S. Amari and S. Wu, "Improving support vector machine classifiers by modifying kernel functions," *Neural Networks*, vol. 12, no. 6, pp. 783–789, Jul. 1999, doi: 10.1016/S0893-6080(99)00032-5.

[161] C. M. Bishop, *Prml*. 2006. Accessed: Mar. 10, 2022. [Online]. Available: https://link.springer.com/book/9780387310732

[162] V. K. Minja Belić, Milica Djurić-Jovičić, Milica Ječmenica Lukić, Igor Petrović, Saša Radovanović, Mirjana Popović, "Implementation of discrete wavelet transformation in repetitive finger tapping analysis for patients with Parkinson's disease," 2016.

[163] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.

[164] A. Jordao, A. C. Nazare Jr, J. Sena, and W. R. Schwartz, "Human Activity Recognition Based on Wearable Sensor Data: A Standardization of the State-of-the-Art," *arXiv preprint arXiv:1806.05226*, 2018.

[165] G. Sergey Ioffe and G. Christian Szegedy, "Batch Normalization," *Icml*, 2015, doi: 10.1007/s13398-014-0173-7.2.

[166] "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." http://proceedings.mlr.press/v37/ioffe15.html (accessed Jul. 14, 2022).

[167] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, 2014, doi: 10.1214/12-AOS1000.

[168] X. Li, S. Chen, X. Hu, and J. Yang, "Understanding the disharmony between dropout and batch normalization by variance shift," *arXiv preprint arXiv:1801.05134*, 2018.

[169] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015. doi: 10.1109/ICCV.2015.123.

[170] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *PMLR*, 2010, doi: 10.1.1.207.2059.

[171] M. Kohlbrenner, T. U. Berlin, R. Hofmann, S. Ahmmed, and Y. Kashef, "Pre-Training CNNs Using Convolutional Autoencoders," Berlin, 2017.

[172] L. Zhou, H. Liu, J. Bae, J. He, D. Samaras, and P. Prasanna, "Self Pre-training with Masked Autoencoders for Medical Image Analysis," Mar. 2022, doi: 10.48550/arxiv.2203.05573.

[173] A. Hartigan and M. A. Wong, "A K-Means Clustering Algorithm," *Journal of the Royal Statistical Society*, vol. 28, no. 1, 1979.

[174] D. P. Kingma and J. Lei Ba, "ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION."

[175] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative Adversarial Networks: An Overview," *IEEE Signal Processing Magazine*, vol. 35, no. 1. 2018. doi: 10.1109/MSP.2017.2765202.

[176]   V. Sandfort, K. Yan, P. J. Pickhardt, and R. M. Summers, "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks," *Scientific Reports 2019 9:1*, vol. 9, no. 1, pp. 1–9, Nov. 2019, doi: 10.1038/s41598-019-52737-x.

[177]   H. C. Shin *et al.*, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11037 LNCS, pp. 1–11, 2018, doi: 10.1007/978-3-030-00536-8_1/COVER/.

[178]   I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein GANs," in *Advances in Neural Information Processing Systems*, 2017, vol. 2017-December.

[179]   J. P. Hart and A. W. Shogan, "Semi-greedy heuristics: An empirical study," *Operations Research Letters*, vol. 6, no. 3, pp. 107–114, Jul. 1987, doi: 10.1016/0167-6377(87)90021-6.

[180]   N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *American Statistician*, vol. 46, no. 3, pp. 175–185, 1992, doi: 10.1080/00031305.1992.10475879.

[181]   K. Marek *et al.*, "The Parkinson Progression Marker Initiative (PPMI)," *Prog Neurobiol*, vol. 95, no. 4, pp. 629–635, Dec. 2011, doi: 10.1016/J.PNEUROBIO.2011.09.005.

[182]   V. Asanza, N. N. Sánchez-Pozo, L. L. Lorente-Leyva, D. H. Peluffo-Ordóñez, F. R. Loayza, and E. Peláez, "Classification of Subjects with Parkinson's Disease using Finger Tapping Dataset," *IFAC-PapersOnLine*, vol. 54, no. 15, pp. 376–381, Jan. 2021, doi: 10.1016/J.IFACOL.2021.10.285.

[183]   M. Mendonça *et al.*, "Inertial sensor-based kinematics in the differential diagnosis of Parkinson's disease and atypical parkinsonisms," *Movement Disorders*, vol. 35, no. SUPPL 1, p. S646, 2020, [Online]. Available: https://www.embase.com/search/results?subaction=viewrecord&id=L633833076&from=export

[184]   J. Song *et al.*, "Differential diagnosis between Parkinson's disease and atypical parkinsonism based on gait and postural instability: Artificial intelligence using an enhanced weight voting ensemble model," *Parkinsonism Relat Disord*, vol. 98, pp. 32–37, May 2022, doi: 10.1016/J.PARKRELDIS.2022.04.003.

[185]   K. Daoudi, B. Das, T. Tykalová, J. Klempir, and J. Rusz, "Speech acoustic indices for differential diagnosis between Parkinson's disease, multiple system atrophy and progressive supranuclear palsy," *NPJ Parkinsons Dis*, 2022, Accessed: Aug. 24, 2022. [Online]. Available: https://hal.inria.fr/hal-03740038

[186]   R. Kowalska-Taczanowska, A. Friedman, and D. Koziorowski, "Parkinson's disease or atypical parkinsonism? The importance of acoustic voice analysis in differential diagnosis of speech disorders," *Brain Behav*, vol. 10, no. 8, 2020, doi: 10.1002/brb3.1700.

[187]   Z.-H. Zhou, "A Brief Introduction to Weakly Supervised Learning," *Natl Sci Rev*, 2017, doi: 10.1093/nsr/nwx106.

[188]   "Sifted - Seed rounds june investment ." https://sifted.eu/articles/seed-rounds-june-investment/ (accessed Aug. 24, 2022).

[189]   B. Beaulieu-Jones *et al.*, "Trends and Focus of Machine Learning Applications for Health Research," *JAMA Netw Open*, vol. 2, no. 10, 2019, doi: 10.1001/jamanetworkopen.2019.14051.

[190] D. Mood *et al.*, "Hamilton M: A rating scale for depression," *J Neurol Neurosurg Psychiatry*, vol. 23, 1960.

[191] A. T. Beck, C. H. Ward, M. Mendelson, J. Mock, and J. Erbaugh, "An Inventory for Measuring Depression," *Arch Gen Psychiatry*, vol. 4, no. 6, pp. 561–571, Jun. 1961, doi: 10.1001/ARCHPSYC.1961.01710120031004.

[192] M. HAMILTON, "THE ASSESSMENT OF ANXIETY STATES BY RATING," *British Journal of Medical Psychology*, vol. 32, no. 1, pp. 50–55, Mar. 1959, doi: 10.1111/J.2044-8341.1959.TB00467.X.

[193] S. E. Starkstein, H. S. Mayberg, T. J. Preziosi, P. Andrezejewski, R. Leiguarda, and R. G. Robinson, "Reliability, validity, and clinical correlates of apathy in Parkinson's disease," *https://doi.org/10.1176/jnp.4.2.134*, vol. 4, no. 2, pp. 134–139, Apr. 2006, doi: 10.1176/JNP.4.2.134.

[194] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "'Mini-mental state': A practical method for grading the cognitive state of patients for the clinician," *J Psychiatr Res*, vol. 12, no. 3, pp. 189–198, Nov. 1975, doi: 10.1016/0022-3956(75)90026-6.

[195] E. Mioshi, K. Dawson, J. Mitchell, R. Arnold, and J. R. Hodges, "The Addenbrooke's Cognitive Examination revised (ACE-R): A brief cognitive test battery for dementia screening," *Int J Geriatr Psychiatry*, vol. 21, no. 11, pp. 1078–1085, Nov. 2006, doi: 10.1002/GPS.1610.

[196] S. Radovanović, M. Jovičić, N. P. Marić, and V. Kostić, "Gait characteristics in patients with major depression performing cognitive and motor tasks while walking," *Psychiatry Res*, vol. 217, no. 1–2, pp. 39–46, Jun. 2014, doi: 10.1016/J.PSYCHRES.2014.02.001.

[197] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity," *J R Soc Interface*, vol. 8, no. 59, pp. 842–855, Jun. 2011, doi: 10.1098/RSIF.2010.0456.

[198] Y. Dupuis, X. Savatier, and P. Vasseur, "Feature subset selection applied to model-free gait recognition," *Image Vis Comput*, vol. 31, no. 8, 2013, doi: 10.1016/j.imavis.2013.04.001.

[199] L. Toloşi and T. Lengauer, "Classification with correlated features: unreliability of feature ranking and solutions," *Bioinformatics*, vol. 27, no. 14, pp. 1986–1994, Jul. 2011, doi: 10.1093/BIOINFORMATICS/BTR300.

[200] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science (1979)*, vol. 315, no. 5814, pp. 972–976, Feb. 2007, doi: 10.1126/SCIENCE.1136800/SUPPL_FILE/FREY.SOM.PDF.

[201] L. Breiman, "Random Forests," *Machine Learning 2001 45:1*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[202] I. Perunicic-Mladenovic and S. Filipovic, "Proneness to Alcohol use Disorder or Pathological Gambling as Differentially Determined by Early Parental and Personality Factors," *J Gambl Stud*, pp. 1–21, Jan. 2022, doi: 10.1007/S10899-021-10095-2/TABLES/4.

[203] Y. Qi, Z. Bar-Joseph, and J. Klein-Seetharaman, "Evaluation of different biological data and computational classification methods for use in protein interaction prediction," *Proteins: Structure, Function, and Bioinformatics*, vol. 63, no. 3, pp. 490–500, May 2006, doi: 10.1002/PROT.20865.

[204] R. Díaz-Uriarte and S. Alvarez de Andrés, "Gene selection and classification of microarray data using random forest," *BMC Bioinformatics*, vol. 7, no. 1, pp. 1–13, Jan. 2006, doi: 10.1186/1471-2105-7-3/FIGURES/1.

[205] K. L. Lunetta, L. B. Hayward, J. Segal, and P. van Eerdewegh, "Screening large-scale association study data: exploiting interactions using random forests," *BMC Genet*, vol. 5, Dec. 2004, doi: 10.1186/1471-2156-5-32.

[206] M. Grajić, I. Stanković, S. Radovanović, and V. Kostić, "Gait in drug naïve patients with *de novo* Parkinson's disease – altered but symmetric," *http://dx.doi.org/10.1179/1743132815Y.0000000043*, vol. 37, no. 8, pp. 712–716, Aug. 2015, doi: 10.1179/1743132815Y.0000000043.

[207] F. Wahid, R. K. Begg, C. J. Hass, S. Halgamuge, and D. C. Ackland, "Classification of Parkinson's disease gait using spatial-temporal gait features," *IEEE J Biomed Health Inform*, vol. 19, no. 6, pp. 1794–1802, Nov. 2015, doi: 10.1109/JBHI.2015.2450232.

[208] S. A. Chatterjee, "Mediolateral stability during gait in people with Parkinson's disease," Graduate Theses and Dissertations. 11377., 2010.

[209] T. E. Raffegeau *et al.*, "A Meta-Analysis: Parkinson's Disease and Dual-Task Walking," *Parkinsonism Relat Disord*, vol. 62, p. 28, May 2019, doi: 10.1016/J.PARKRELDIS.2018.12.012.

[210] M. Amboni *et al.*, "Gait analysis may distinguish progressive supranuclear palsy and Parkinson disease since the earliest stages," *Sci Rep*, vol. 11, no. 1, p. 9297, Dec. 2021, doi: 10.1038/S41598-021-88877-2.

Author biography

Minja Belić was born on 16th May 1986 in Valjevo, Serbia, where she finished primary school and Valjevo Gymnasium, enrolled in a specialized mathematics class. In 2012 she graduated from the School of Electrical Engineering, department of Physical Electronics, sub-department Biomedical Engineering with a thesis exploring an image processing approach to Iris Recognition, under the guidance of prof. dr Irini Reljin. She obtained her master's degree from the same institution in 2013, with the thesis titled "Surface electrical stimulation of afferent fibers of the forearm for sensory substitution" mentored by prof. dr Dejan Popović. She started her multidisciplinary PhD studies at the University of Belgrade in 2014, at the department of Biomedical engineering and technologies.

From 2013 to 2018 she worked with Tecnalia Serbia, a research and development oriented company, where she took part in several projects aiming to utilize surface electrical stimulation in rehabilitation (stroke rehabilitation), pain management (lower back), slowing the progression of osteoporosis (spine), and sensory substitution (using electrical stimulation to close the feedback loop from myoelectric hand prostheses for trans-radial amputees).

She cooperated with the Innovation center of School of Electrical Engineering and Clinical Center of Serbia, while working on sensors and algorithms for diagnostic aid in Parkinson's disease and atypical parkinsonisms.

From 2019 to 2022 she worked in Novelic on radar solutions for human detection and contactless cardio-respiratory monitoring, with a focus on automotive applications. She worked as an algorithm design engineer and subsequently technical and delivery project manager, and initiated the formation of a specialized machine learning team within the company.

Since 2022 she has been employed as a data scientist in Daon Inc, where she works on image processing and machine learning algorithms for face biometrics.

Minja Belić co-authored 7 scientific papers published in journals on the SCI list, 5 conference papers and 4 papers in national and other journals.

Изјава о ауторству

Име и презиме аутора: Миња Белић

Број индекса: 13/2014

Изјављујем

да је докторска дисертација под насловом „Примена алгоритама вештачке интелигенције за обраду кинематичких сигнала у дијагностици Паркинсонове болести и атипичних паркинсонизама“:

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршила ауторска права и користила интелектуалну својину других лица.

У Београду, __30.3. 2025.__   Потпис аутора

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора: Миња Белић

Број индекса 13/2014

Студијски програм: Биомедицинско инжењерство и техноилогије

Наслов рада: Примена алгоритама вештачке интелигенције за обраду кинематичких сигнала у дијагностици Паркинсонове болести и атипичних паркинсонизама

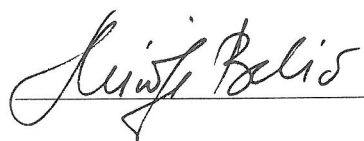Ментори: проф. др Захарије Радивојевић, др Саша Радовановић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у **Дигиталном репозиторијуму Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

У Београду, **30.3.2023.**　　　　　　　　　　Потпис аутора

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић" да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

„Примена алгоритама вештачке интелигенције за обраду кинематичких сигнала у дијагностици Паркинсонове болести и атипичних паркинсонизама"

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучила.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци. Кратак опис лиценци је саставни део ове изјаве).

У Београду, 30. 3. 2025.   Потпис аутора

1. **Ауторство.** Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално.** Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада.** Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима.** Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада.** Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима.** Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.