Multi-Modal Deep Learning Diagnosis of Parkinson's Disease—A Systematic Review

Vasileios Skaramagkas, Anastasia Pentari, Zinovia Kefalopoulou, and Manolis Tsiknakis

Abstract—**Parkinson's Disease (PD) is among the most frequent neurological disorders. Approaches that employ artificial intelligence and notably deep learning, have been extensively embraced with promising outcomes. This study dispenses an exhaustive review between 2016 and January 2023 on deep learning techniques used in the prognosis and evolution of symptoms and characteristics of the disease based on gait, upper limb movement, speech and facial expression-related information as well as the fusion of more than one of the aforementioned modalities. The search resulted in the selection of 87 original research publications, of which we have summarized the relevant information regarding the utilized learning and development process, demographic information, primary outcomes, and sensory equipment related information. Various deep learning algorithms and frameworks have attained state-of-the-art performance in many PD-related tasks by outperforming conventional machine learning approaches, according to the research reviewed. In the meanwhile, we identify significant drawbacks in the existing research, including a lack of data availability and interpretability of models. The fast advancements in deep learning and the rise in accessible data provide the opportunity to address these difficulties in the near future and for the broad application of this technology in clinical settings.**

Index Terms—**Artificial intelligence, deep learning, deep neural networks, Parkinson's disease, speech, facial expressions, gait, upper limbs, tremor.**

I. INTRODUCTION

PARKINSON'S disease (PD) is the second most common neurological disease after Alzheimer's, affecting 2-3% of the population over 65 years of age globally [1]. Neuronal loss in the substantia nigra, which causes dopamine deficiency in the striatum, and intracellular inclusions containing aggregates of α-synuclein are the neuropathological hallmarks of PD. Moreover, environmental and lifestyle factors, sex and age,

Vasileios Skaramagkas and Manolis Tsiknakis are with the Biomedical Informatics and eHealth Laboratory, Department of Electrical and Computer Engineering, Hellenic Mediterranean University, Heraklion, 71410 Crete, Greece, and also with the Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH), Heraklion, 70013 Crete, Greece (e-mail: vskaramag@ics.forth.gr).

Anastasia Pentari is with the Institute of Computer Science, Foundation for Research and Technology Hellas (FORTH), Heraklion, 70013 Crete, Greece.

Zinovia Kefalopoulou is with the Department of Neurology, Patras University Hospital, 26404 Patras, Greece.

Digital Object Identifier 10.1109/TNSRE.2023.3277749

in combination with genetics, all contribute to the pathogenesis of PD.

Even though there is presently no cure for PD, pharmacological approaches based on dopamine substitution, as well as surgical techniques such as deep brain stimulation (DBS) provide substantial improvement of motor symptoms. While mortality is not increased in the first decade after disease onset, eventually doubles compared to the general population, with falls and aspiration pneumonia being the leading causes of hospitalization and decease.

PD is clinically defined by the presence of bradykinesia and at least one additional cardinal motor symptom (rigidity or rest tremor), in combination with other supporting and exclusionary features [2], [3]. Reduced facial expression, handwriting diminishing (micrographia), speech and voice impairment, as well as difficulty swallowing, all represent additional motor aspects of PD. The aforementioned motor symptoms gradually interfere with patients' daily activities, negatively affecting their quality of life. Progression of motor dysfunction with increasing gait abnormalities and onset of postural instability further compromise patients' autonomy and safety. Although the motor symptoms define the clinical syndrome, a majority of PD patients have other complaints that have been classified as non-motor, probably related to non-dopaminergic pathways. These include mood and mental alterations, such as depression, lack of motivation or apathy, and a declining cognitive capacity. Fatigue, sleep disturbances, autonomic (i.e., orthostatic hypotension, urogenital dysfunction, constipation, and excessive sweating), as well as sensory complaints are common components of the clinical spectrum of PD. Non-motor symptoms are present in the early stages of the disease, while some antedate the onset of cardinal motor features by years even decades [4]. However, these symptoms become increasingly prevalent and are major determinants of quality of life, progression of overall disability and mortality, as the disease advances [5].

The diagnosis of Parkinson's disease is mostly dependent on the patient's clinical evaluation [6], [7]. The Movement Disorder Society - Unified Parkinson's Disease Rating Scale (MDS-UPDRS) is currently regarded as the gold standard for the evaluation and monitoring of PD [8], followed by the well-established Hoehn and Yahr (H&Y) scale. Although there are some interesting possibilities, there are no clear biomarkers for PD, and all pertinent research findings support the clinical diagnostic confirmation. Even more difficult is

the identification of the disease in its early stages, since the symptoms' heterogeneous appearance and progression result in complicated clinical presentations of PD [9]. Fortunately, Machine Learning (ML), a field of Artificial Intelligence (AI) which is defined as the capability of systems to autonomously acquire knowledge and detect patterns from experience or existing data without being explicitly programmed [10], has become increasingly successful in identifying nonlinear connections in high-dimensional data. Moreover, a cutting-edge machine learning technology, Deep Learning (DL), has recently achieved success and increased performance that surpasses state-of-the-art in several health areas [11]. DL allows the entry of high-dimensional and unprocessed data and automatically learns its representation through the use of Deep Neural Networks (DNN), which need minimum feature engineering work on data preprocessing [12].

DL strategies are essential to manage such complex data extracted from sensor-enabled devices [13]. During the last decade, there has been a rapid development in the use of DL modeling techniques that employ sensor data to track, monitor, and predict the course of Parkinson's disease [14], [15]. Other techniques involve the use of biomarker data sets, such as dopamine transporter data from tomographic images and serum cytokines, to evaluate classification performance using shape features derived from produced regions of interest [16]. Additionally, a recent study provided evidence that AI can identify people, who suffer from PD, from their nocturnal breathing and could accurately estimate the disease severity and progression [17].

Nowadays, research on the development of wearable sensory equipment for the detection of symptoms or characteristics of PD has made significant progress [18]. The sensors can be located in different parts of the body and are associated with different modalities and combinations thereof. The proper use of sensors and the selection of the investigated modalities can offer provision to more remote healthcare services and tailored diagnoses [19]. In addition, these digital sensors enable the autonomous, non-disruptive collection of real-world data [20]. Meanwhile, the real-time nature of wearable technology and the depth of insight into a patient's vitals enable physicians to detect disease early and provide the most accurate diagnoses, whether in-clinic or remotely [19]. Therefore, it is of utmost importance the use of sensors to obtain biomarkers from modalities, that can offer real-time monitoring and thus, may accurately reflect everyday symptoms and their variation.

In the presented systematic literature review (SLR) we focus on studies that employ DL methods towards PD diagnosis by utilizing data from the modalities of gait, upper limb motion, speech and facial expressions. Information from the aforementioned four modalities can be obtained from wearable devices during daily life activities in real-time, outside of the clinical settings, and thus, can provide caregivers and clinicians with a more realistic picture of the progression of the disease, leading to more personalized treatment of the patient.

The rest of the manuscript is structured as follows: Section II provides with a description of the most common PD symptoms related to the investigated modalities and the importance of their use for PD diagnosis is highlighted.

Further, Section III comprises of the related recently published reviews regarding the use of AI to the diagnosis of PD. Further, the research method followed in this review is presented in detail in Section IV, including the research goal and questions, the study search and selection strategy as well as the extraction of information from the selected surveys. In Section V, our findings are analyzed and the research questions are answered in detail. Finally, in Section VI, we summarize and discuss our results and the possible limitations of our work whilst in Section VII we draw the main conclusions from our review.

II. Manifestation of PD Symptoms

In this Section, the clinical features of Parkinson's disease regarding the modalities of gait, upper limb motion, speech and facial expressivity, both motor and non-motor, are described in the context of the progression of the disease.

A. Analysis of Gait

Gait impairment is an evolving condition, and different patterns of gait disturbances can be detected throughout the progression of PD, such as reduced smoothness of locomotion, increased interlimb asymmetry [21], decreased speed, reduced step length [22], shuffling steps, increased double-limb support, defragmentation of turns, decreased balance and postural control [23].

In addition, data related to gait mechanisms are utilized to create strategies for predicting the risk of falling. Falls represent a significant risk factor that greatly influences the quality of life of Parkinson's patients [24]. Falls in people with severe Parkinson's disease are often related to a paroxysmal symptom known as "Freezing of Gait" (FoG). More than 60% of PD patients experience some form of freezing as the disease progresses, and insufficient knowledge of the pathophysiology and circuit mechanisms limits the effectiveness of treatment [25]. Considering this, FoG is characterized as an episodic (seconds-long) failure to create an effective step, and its pathophysiology likely involves context-dependant dysfunction across multiple neuronal levels, including cortical, subcortical and brainstem regions [26]. FoG can occur at the beginning of the first step, when turning, performing several tasks, walking through narrow spaces, arriving at a destination, or walking through doors. Accelerated pacing (festination), another sign of Parkinson's disease, is clinically characterized as the patient's tendency to walk forward in increasingly rapid and shorter steps, with the patient's center of gravity shifting forward, over the supporting leg. There is evidence that individuals with Parkinson's disease who frequently present FoG follow a pattern of accelerated pacing prior to "freezing" [27]. Several studies show that, among a number of gait metrics, step duration, step length, double support time and swing time might be used to create models that predict episodes of FoG [28].

Therefore, an objective and quantitative analysis of the modality of gait could potentially improve the current approach, which may aid in Parkinson's disease patient diagnosis, symptom monitoring, therapy management, rehabilitation, and fall risk assessment and prevention.

B. Analysis of Upper Limb Motion

Typically, the motor manifestations of PD have a localized onset (i.e. reduced arm swing, hand tremor [29]). Uncertainty exists over the appearance of a recognized somatotopic pattern of development and progression of motor symptoms in PD. According to patterns based on postmortem studies and in vivo positron emission tomography (PET) imaging, motor indications of Parkinson's disease should begin in the lower extremities and progress upward [29]. However, the majority of neurologists would concur that the early motor symptoms of PD manifest in the upper limbs. This is a subjective conclusion based on the experience, which might be influenced by observational bias. Therefore, the motor manifestations of the upper limb may be more apparent to patients and clinicians than the early, modest impairments of the lower limb and face [30].

Tremor, which is defined as an involuntary, rhythmic, and oscillating movement of a body part, is one of the most common symptoms of PD [31] and is referred as Parkinsonian tremor (PT). The objective of the tremor analysis is to quantify PD tremor, which is rhythmic, has a typical frequency, and is more prevalent in the hands during rest [32]. There are various types of tremor, each with unique causes and characteristics [33]. The clinical picture, the correct interpretation of medical history, and the doctor's experience all contribute to a more precise determination of tremor type. Despite the existence of a variety of diagnostic methods for Parkinson's disease and tremor, their rapid and effective differentiation, especially in their early stages, proves especially challenging due to their wide variety of causes and symptom similarities [34].

In addition to tremor episodes, freezing phenomena are a significant cause of disease-related disability in PD [35]. Freezing occurs most frequently during walking (FoG), as well as during swallowing, speech, and especially repetitive movements of the upper limbs [36]. As with FoG, freeze of the upper limbs impairs daily activities such as handwriting, tooth brushing, typing, and bimanual coordination to a significant degree [37]. Consequently, it is essential to determine the cause and detect the abnormalities in the motion of the upper limbs in order to effectively reduce and treat the associated syndrome.

C. Analysis of Speech

Speech is an essential biological feature of human beings and impairments in voice can indicate possible disorders, such as PD. The majority of PD patients develop a variety of speech disorders while speech might be affected years before the main motor symptoms of the disease appear [38]. Therefore, the modality of speech is crucial to be investigated early and in depth.

The most important and common impairments include reduced loudness, monopitch, monoloudness, breathy and hoarse voice quality, reduced stress and imprecise articulation [39], [40]. These changes in speech belong to the hypokinetic dysarthria category and can be exploited as possible biomarkers so as to identify early signs of this neurological disorder [41]. Sustained vowels as well as reading sentences, words and short texts have been employed to distinguish the PD patients from healthy controls [38], [42]. Moreover, imprecise articulation, phonation and prosody have been observed in specific consonants, as mentioned in [41]. All these symptoms can be characterized as mild, moderate or severe [38].

D. Analysis of Facial Expressions

Facial expressions constitute a fundamental source of information for disease analysis, manifested in the early stages of the disease, even years before diagnosis [43]. Hypomimia, also known as facial amimia, is one of the most prominent clinical indications of Parkinson's disease, characterized by a decrease or lack of spontaneous facial movements, small-amplitude and low-velocity voluntary orofacial movements, and emotional expressiveness. Moreover, it has been reported that there is a possible relationship between amimia and other axial symptoms, such as gait freezing [44].

Observable signs include, among others, abnormalities associated with a deeper indentation of the eyelids, a staring expression, involuntary mouth opening, and stiffness in the orbicularis oculi muscles [45]. On the upper face, hypomimia often is often presented as a reduction in blink rate, but on the lower face, spontaneous smiling difficulties are noticed [46]. In addition, the early manifestations of facial muscle movement-related symptoms along with their possible correlation with motor symptoms of PD highlight the need for more research to be conducted on the alterations of facial expressions.

III. RELATED WORKS

In recent years, the number of publications on the application of deep learning to the diagnosis of PD has increased. Although previous studies have reviewed the use of machine learning in the diagnosis and assessment of PD, they were limited to the analysis of motor symptoms, kinematics, and wearable sensor data as well as the utilization of traditional ML techniques [47], [48]). Recently, [49] provided a comprehensive analysis of the influence of machine learning and deep learning approaches applied towards Parkinson's disease diagnosis on the development of new research areas based on neuroimaging techniques as well as physiological signals obtained from speech, gait and handwriting. In addition, this research investigates the existing status and potential applications of data-driven AI technologies in the diagnosis of Parkinson's disease.

Healthcare services are gaining interest in computer-aided diagnostic (CAD) technologies based on artificial intelligence methods that can perform automated diagnosis of Parkinson's disease. To this goal, authors in [50], collected 63 studies (2011-2021) on deep learning from various modalities including brain analyses (SPECT, PET, MRI, and EEG) and motion symptoms (gait, handwriting, speech, EMG). They demonstrated that deep learning models can reach excellent prediction accuracy for Parkinson's disease, particularly the convolutional neural network (CNN) model, which has been widely recommended by studies focusing on image classification for brain imaging and handwriting analysis. Additionally,

the CNN model worked well with one-dimensional inputs such as EEG and speech analysis. This work suggests that academics will be incentivized to use more explainable and interpretable methodologies in deep learning-based CAD tools, which will subsequently be adopted by end-users and enhance the health care results for the increasing number of people affected by PD around the world.

In order to offer a complete overview of the data modalities and machine learning algorithms utilized in the diagnosis and differential diagnosis of PD, [51] conducted a literature evaluation of papers published through 2020, using PubMed and IEEE Xplore. This study examines the aims, data sources and types, machine learning methods, and associated outcomes of 209 papers that were included. These results reveal a strong potential for the implementation of ML techniques and novel biomarkers in clinical decision making, resulting in a more systematic and accurate diagnosis of Parkinson's disease.

In this systematic review, we aim to (a) comprehensively summarize published studies that applied advanced DL models to the diagnosis of PD-related symptoms and estimation of disease severity levels for an exhaustive overview of data sources, data types, deep learning models, and associated outcomes, (b) assess and compare the feasibility and efficiency of the different DL methods in the diagnosis of PD by utilizing information obtained from four different modalities: gait, upper limb motion, speech, and facial expressions, (c) provide machine learning practitioners interested in the diagnosis of PD with an overview of previously used models and data modalities and the associated outcomes as well as the different PD diagnostic targets in relation to each modality studied, and (d) present and comment upon the variety and types of sensors used to acquire the desired information from the studies modalities. The application of AI to clinical and non-clinical data of different modalities has often led to high diagnostic accuracies in human participants, therefore may encourage the adaptation of cutting-edge DL algorithms and novel biomarkers in clinical settings to assist more accurate and informed decision making.

IV. Research Method

This study was designed and carried out by following guidelines for systematic literature reviews [52]. We followed the process depicted in Fig. 1, which can be divided into the three common phases of planning, conducting, and documenting.

The objective of Phase 1, i.e. the Planning phase, was to:
- establish the need for a review of DL methods implemented towards the diagnosis of PD.
- identify research goal and more importantly questions, and
- define the protocol to be followed by the research team for carrying out the work in a systematic and pinpoint manner.

The output of the Phase 1 is a detailed review protocol.

The objective of Phase 2 was to perform the SLR by carrying out all the steps defined in the review protocol, as follows:
- Search and selection: Four peer-reviewed databases were searched automatically. Then, possible entries were

filtered to create the final list of studies to be reviewed. After selection, we performed comprehensive backward and forward snowballing.
- Data extraction form definition and classification framework: We compared and categorized studies considering research topics [53]. This was accomplished methodically with keywording.
- Data extraction and synthesis: We examined each main research in depth, hence completing the associated data extraction form. Forms were gathered and aggregated for further analysis and synthesis. We also reviewed and analyzed the previous data. This task elaborated on extracted data to address each research question.

In Phase 3, we analyzed and synthesized the data. For independent replication and verification, we developed a thorough replication kit. We include search and selection raw data, the full list of important research, and data extraction raw data.

A. Research Goal and Questions

The goal of this study was to identify, classify, and evaluate trends, focus, and open challenges in existing research on the diagnosis of Parkinson's disease by utilizing cutting-edge deep learning techniques. Our focus is to search for deep learning methods utilized towards the detection of various indicators of Parkinson's disease such as motor and non-motor symptoms, severity levels or other PD related characteristics [54].

Our overall goal can be refined into the following specific research questions (RQ), for each of which we also provide primary objective of investigation:

- **RQ1**: *What are the most advanced DL methods used for the diagnosis of PD and the assessment of PD severity using the modalities of interest?*
- **RQ2**: *Has DL being used for non-conventional PD-related diagnostic targets?*
- **RQ3**: *Could the fusion of multi-modal information offer more personalized and accurate diagnosis?*
- **RQ4**: *Which sensing systems are used to analyze the symptoms of PD related to speech, facial expressions, upper limb movement and gait?*

Based on the review results obtained by answering RQ1, we offer a deep knowledge of the most advanced DL techniques and frameworks employed towards the diagnosis of PD, its symptoms and severity levels from data obtained from the modalities of speech, gait, upper limb motions and facial expressiveness. Answering RQ2, gives an insight regarding the existence of DL models aiming to the identification of non-conventional PD diagnostic targets. Moreover, answering RQ3 helps the community to understand whether there is potential in the adoption of DL methods based on the fusion of different modalities and space for improvement in this research area. Finally, by answering RQ4 we provide a solid foundation for a thorough comparison of existing sensory equipment solutions that are currently used to extract the necessary information which are then processed and served as input to the DL models.

contribution is useful for both (i) researchers to further contribute to this research area by defining new approaches
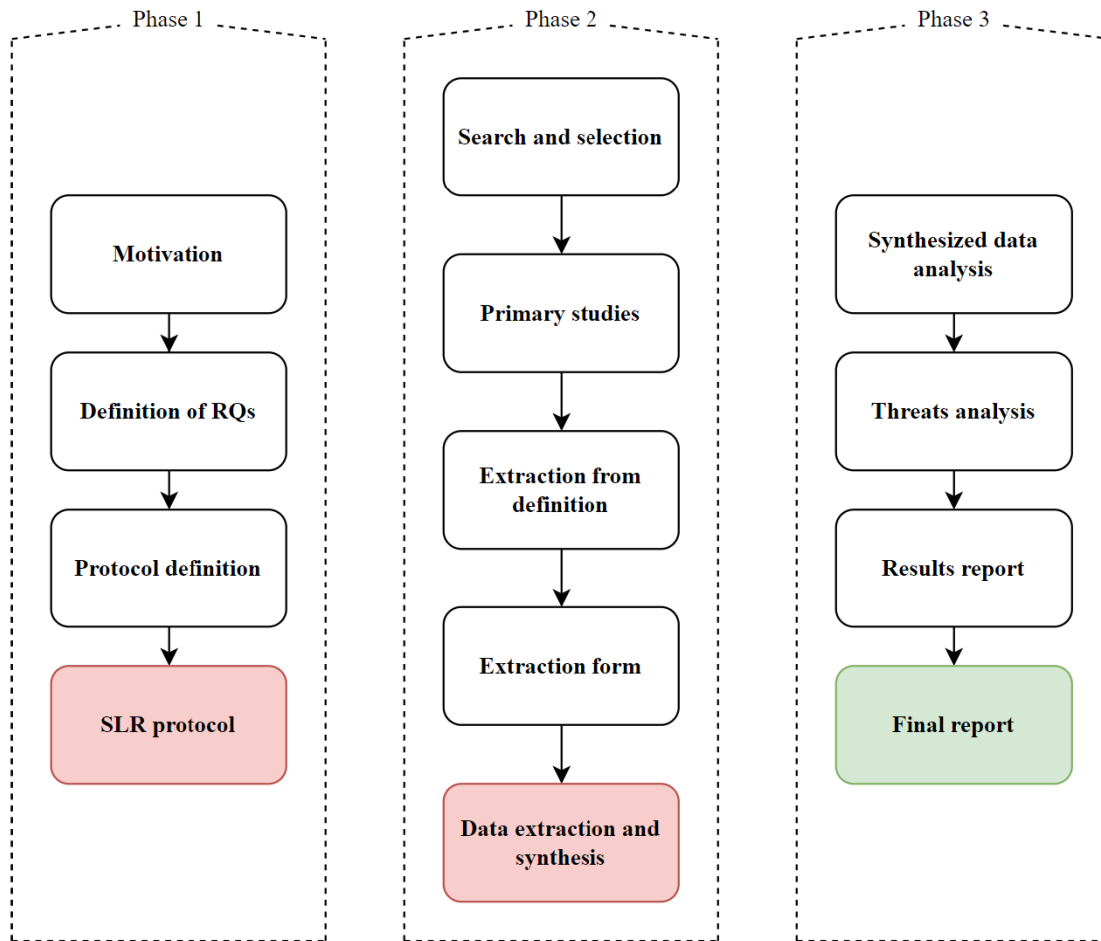
Fig. 1. Overview of the SLR process.

or refining existing ones, and (ii) practitioners to better understand existing methods and techniques and thereby to be able to adopt the one that better suites their research and business goals.

B. Search and Selection Strategy

In this phase, we gathered the set of research studies that are relevant and representative for our purposes. Before performing the actual search and selection of relevant studies, we manually selected a set of ten pilot studies. They were selected based on the authors' knowledge of the targeted research domain (i.e., DL methods for multi-modal PD diagnosis) and on an informal preliminary screening that we performed on the available literature on the topic. Selected pilot studies fulfil our selection criteria (see below) and they are presented in Table I. Pilot studies were used to validate our search and selection strategy; more specifically, we used them to have quick feedback about the goodness of our search string to be used for the automatic search and for guiding the refinement of the selection criteria.

1) Automatic Search: During this phase, automated searches were conducted on the electronic databases and indexing systems specified in Table II. As indicated in [64], we selected four of the largest and most comprehensive scientific databases and indexing systems in biomedical engineering, namely SCOPUS, IEEE Xplore Digital Library, PubMed, and ACM Digital Library in order to cover as much relevant material as possible. The selection of these electronic databases and indexing systems was strongly affected by their high accessibility, their capacity to export search results to well-defined, computation-friendly formats, and the fact that they have been acknowledged as an efficient way to conduct systematic literature reviews in biomedical engineering.

To create the search string, we considered initially the research questions and afterwards the set of pilot studies. Then, we retrieved a list of relevant concepts, their synonyms, abbreviations, and alternative spellings, and merged them with ANDs and ORs to form the final search string. The search string displayed below was evaluated by conducting pilot searches on the four sites and comparing the findings to all pilot studies, which were required to be included in the acquired results. The actual search strings used for each database were obtained by syntactically modifying them to the unique database's properties. We searched for the search string on the paper's title, abstract, and keywords; the automated searches yielded 3342 possible studies.

<div align="center">

parkinson*

AND

</div>

TABLE I
PILOT RESEARCH STUDIES

Title	Year	Ref.
Multi-Scale Sparse Graph Convolutional Network For the Assessment of Parkinsonian Gait	2022	[55]
Detecting Freezing of Gait in Parkinson's Disease Patient via Deep Residual Network	2021	[56]
CNN-Based PD Hand Tremor Detection Using Inertial Sensors	2021	[57]
Evaluation of Recurrent Neural Network Models for Parkinson's Disease Classification Using Drawing Data	2021	[58]
Automated video-based assessment of facial bradykinesia in de-novo Parkinson's disease	2022	[59]
Improving Parkinson Detection using Dynamic Features from Evoked Expressions in Video	2021	[60]
Deep Learning-Based Parkinson's Disease Classification Using Vocal Feature Sets	2019	[61]
Parkinson Disease Detection Based On Speech Using Various Machine Learning Models and Deep Learning Models	2021	[62]
End-2-end modeling of speech and gait from patients with parkinson's disease: comparison between high quality vs. smartphone data	2021	[63]
Biomechanical parameters assessment for the classification of Parkinson Disease using Bidirectional Long Short-Term Memory	2020	[64]

TABLE II
ELECTRONIC DATABASES AND INDEXING SYSTEMS
CONSIDERED IN THIS RESEARCH

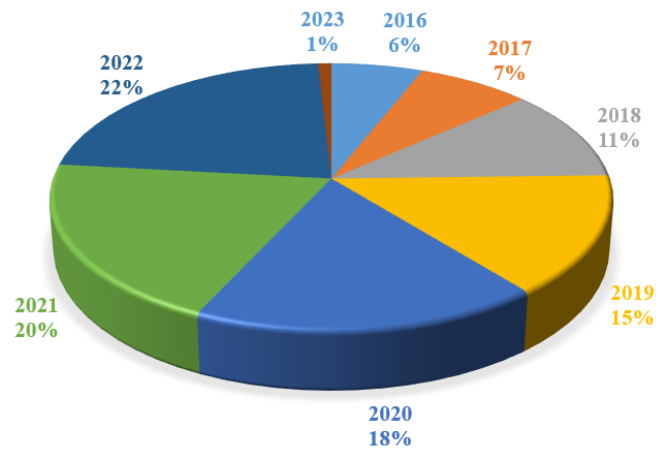| Name	Type	Findings
SCOPUS	Indexing system	1476
IEEE Xplore Digital Library	Electronic database	462
PubMed	Electronic database	433
ACM Digital Library	Electronic database	971

Fig. 2. Percentage of papers included between January 2016 and January 2023.

disease OR phenotyp* OR symptom* OR stage* OR severity
AND
diagnos* OR assess* OR identif* OR classif* OR recogn*
AND
lower limb* OR gait OR face OR facial OR speech OR voice OR tremor OR upper limb* OR multimodal* OR multi-modal*
AND
machine learning OR deep learning OR neural net*

2) Impurity and Duplicates Removal: Due to the nature of electronic databases and indexing systems, search results may contain items that are obviously not research papers, such as conference and workshop proceedings, international standards, textbooks, book chapters, etc., as well as duplicates. At this step, we manually eliminated impurities and merged duplicates. Fig. 2 presents the percentage of the studies included between years 2016 and 2023.

3) Selection Criteria: After removing impurities and duplicates, our inclusion and exclusion criteria were applied to all the remaining research to determine their possible inclusion in the set of studies. Each study was analyzed in two steps: initially, its title, keywords, and abstract were considered; secondly, if the analysis did not result in a clear conclusion, the introduction and conclusion parts were reviewed. The following selection criteria were applied:

Inclusion Criteria for Peer-Reviewed Literature:

1) Studies proposing DL-based PD diagnostic methods based on gait, speech, facial and upper limb movement characteristics.

2) Studies subject to peer review [65].
3) Studies written in English.
4) Studies available as full-text.

Exclusion Criteria for Peer-reviewed literature

1) Secondary and tertiary studies (e.g., systematic literature reviews, surveys).
2) Studies in the form of tutorial papers, short papers (≤ 3 pages), poster papers, editorials, manuals, because they do not provide enough information.
3) Studies that do not involve subjects diagnosed with PD, but healthy volunteers mimicing movements and symptoms caused by the disease.
4) Studies that do not include any information about the DL models that were employed.

To select studies objectively, two researchers actively participated in this phase (V.S. and A.P.). More specifically, by following the method proposed in [66], each potentially relevant study was classified by the researchers as relevant, uncertain, or irrelevant according to the selection criteria above. Studies classified as irrelevant by both raters were immediately excluded, while those marked as relevant were preliminary included. For the uncertain cases, the selection team discussed with the mediation of one more researcher, the
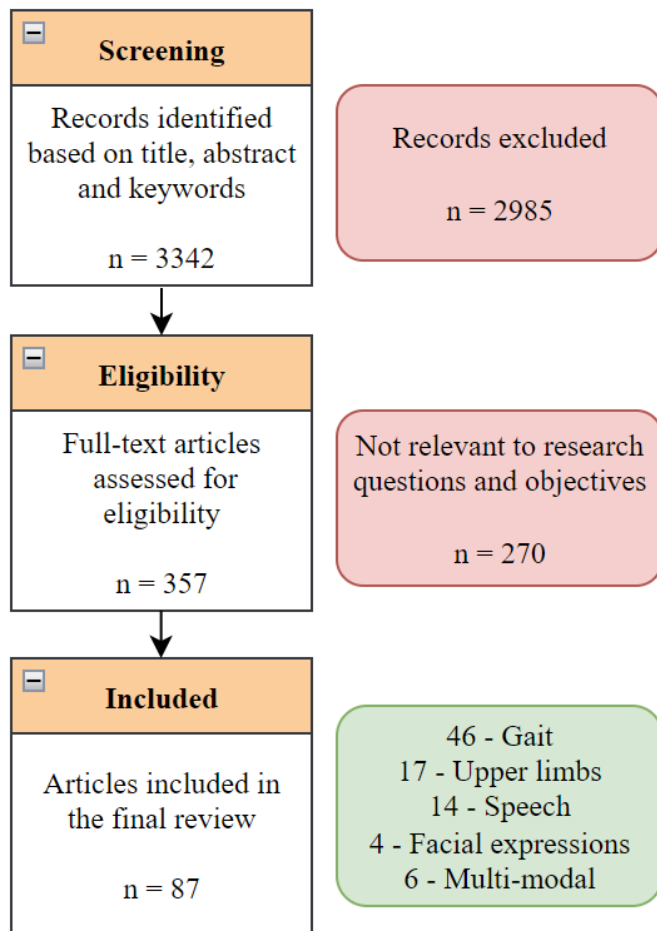
Fig. 3. Prisma model depicting no. of records included and excluded.

TABLE III
DATA EXTRACTION FORM AND ITS CATEGORIES

Category	Description
References	list of contributed authors in each publication
Subjects	The number of subjects participated
Experiment(s)	Followed procedure for data gathering
Diagnosis	Diagnosis target
Development process	Data pre and post-processing, filtering and preparation of input to the model
Model	DL models and methods employed
Performance metrics	Outcome of the models based on testing data

necessary, modified the data extraction form to better accommodate the data; previously analyzed studies were re-analyzed using the modified data extraction form. This procedure was not complete until all studies had been analyzed. The final total number of studies considered and evaluated was 87.

V. DATA ANALYSIS AND SYNTHESIS

In this Section we will focus in providing answers to the questions framed in Section IV, based on the analysis of the available evidence in the studies reviewed. Each subSection will cover one distinct research question.

A. Advanced DL Methods for PD Diagnosis and Severity Assessment Based on the Investigated Modalities

In the next subsections we will present the advanced deep learning techniques and frameworks that have been reported in the literature for the detection of PD's intensity and evolution based on the analysis of gait, upper limb motion, speech and facial expressiveness. With the term "advanced" we refer to recent DL techniques that go beyond the most commonly used methods and either are based on the fusion of DL algorithms or either explore cutting-edge techniques such as attention models, autoencoders, and Generative Adversarial Networks (GANs) [67], [68]. The presented literature review identified a number of advanced DL-based models built for PD diagnosis based on the discussed four modalities. The results are depicted in Tables IV - V.

1) Analysis of Gait: As described in Section II, the objective and quantitative analysis of the modality of gait could potentially lead to a more automatic diagnosis of PD and its symptoms. Based on the findings of the presented literature review, advanced DL methods have been developed accepting as input gait signals during various walking trials. Notably, the majority of studies aim towards the automatic identification of FoG-related events, as until now, the assessment of FoG requires well-trained experts to perform time-consuming annotations via real-time or vision-based observations. The respective studies are presented in Table IV.

In order to make a clinical diagnosis, it is time-consuming and subjective for professionals to evaluate the patient's stride. Currently, the formulation of FoG identification as a human action recognition job in video analysis offers a viable answer

mediator (MT). The Prisma model, shown in Fig. 3, depicts the overall paper review process, the studies included, and provides the number of research papers involved at each stage of the pipeline.

C. Data Extraction

At the beginning of this phase, we have developed a data extraction form to be used to collect data retrieved from each primary study, as shown in Table III. For addressing specific questions about the identified research, we took into account standard information such as the title, authors, type, and publication year of each study. For the research questions, we followed a systematic procedure based on keywording for defining the characteristics of each cluster of the data extraction form and obtaining the corresponding data from the studies.

The objective of the keywording was to create an extraction form that was compatible with previous research and took their features into consideration [53]. Specifically, we gathered keywords and topics by reviewing the complete texts of pilot studies. The gathered keywords and concepts were then clustered in order to organize them according to the selected categories. During the actual extraction process, we obtained any relevant information that did not fit on the data extraction form. We assessed the collected additional data and, when

to these problems. Nevertheless, the majority of existing human action detection algorithms are inadequate for this task, as FoG is extremely delicate and is readily disregarded when it is obscured by irrelevant motion. A novel action detection technique, the convolutional 3D attention network (C3DAN), comprised of a Spatial Attention Network (SAN) and a 3-dimensional convolutional network (C3D), was developed to address this issue [69]. SAN seek to produce a coarse-to-fine attention area, whereas C3D extracted discriminative characteristics. The suggested method can pinpoint the attention region without requiring manual annotation and extract discriminative characteristics from beginning to conclusion. The suggested C3DAN approach for quantifying FoG in PD was assessed on a video dataset gathered from 45 PD patients in a clinical environment and achieved an accuracy of 79.30%. Further, acceleration signals consisted of 237 FoG occurrences were acquired from 10 PD patients' lower back during walking trials in [70]. From successive gait cycles, acceleration patterns and spectrograms were generated and used for FoG detection model training and various domain contributions to FoG detection model training were evaluated by comparing the model's performance. Deep convolutional layers in conjunction with recurrent layers (DeepCNN-LSTM) were learned offline and then used to detect FoG or non-FoG events. The authors achieved an average FoG detection accuracy of 94.30% using the acceleration spectrogram as input.

Another research attempt [71] combined CNN and LSTM networks for FoG recognition. In particular, the authors proposed a FoG detection system in which hand-crafted features were fed to a hybrid CNN-LSTM model for additional feature learning and classification. The manually produced features with time-frequency representation were recovered from the raw sensor data using a discrete wavelet transform (DWT) on many levels. The CNN and bidirectional long short-term memory network (BiLSTM) hybrid deep learning architecture was then utilized to extract deep features and categorize FoG events with an accuracy of 90.01%. Concerning hybrid neural networks, researchers in [72] outlined a deep learning-based strategy consisting of a hybrid Neural Network constructed by merging a CNN, LSTM, and DNN. This approach produced a classification accuracy that was 3.90% higher than comparable efforts. Finally, squeezed and excited deep learning was utilized to solve the FoG detection issue using wearable sensors in [73] by proposing a SE-CNN deep learning model. Each convolutional layer integrated channel-specific input from the squeeze and excitation module due to this mechanism. The Daphnet dataset was utilized to assess the suggested deep learning model, which obtained 95.66% accuracy when compared to other based methods.

Since FoG events may be identified through the motion patterns of joints, scientists defined vision-based FoG detection as a fine-grained graph sequence modelling challenge by modeling the anatomic joints in each temporal segment with a directed graph [74]. To describe FoG patterns, a novel graph sequence recurrent neural network (GS-RNN) with graph recurrent cells that accept graph sequences of dynamic structures as inputs was presented. Experimental results on more than 150 films collected from 45 patients indicated that the suggested GS-RNN for FoG identification displayed good performance with an AUC of 0.80. To overcome the lack of data, patient-independent models were employed to identify FoG and demonstrated high sensitivity but low specificity, or vice versa. Authors in [75] created a Deep Gait Anomaly Detector (DGAD) employing a transfer learning-based technique to increase FoG detection accuracy, while examining the influence of data augmentation and extra pre-FoG segments on the prediction rate. Seven patients with Parkinson's disease undertook a variety of everyday walking exercises while wearing inertial measuring units with the target models accounting for 87.40% of FoG onsets. Additionally, researchers in [76], presented the classification of FoG episodes utilizing Wi-Fi and radar imaging by leveraging multiresolution scalograms formed by channel state information (CSI) imprint and micro-Doppler signatures produced by reflected radar signal. 120 participants participated in a variety of activities, including walking at varying speeds, voluntary stops, sitting and standing, and FoG-inducing exercises. Combining the pictures received from both sensing approaches, the suggested improved Autoencoder was utilized to identify FoG episodes via a data fusion procedure. Using data fusion, the suggested technique achieved an overall accuracy of 98.00%.

Due to the limited number of patients and characteristics in the most frequent utilized datasets for the use of deep learning, the bulk of research concentrate on binary classification tasks, particularly the distinction between PD and non-PD subjects and the presence of FoG in PD patients. However, the early detection of varying degrees of disease severity would additionally aid doctors and likely result in a more individualized disease evaluation.

Following this perspective, a novel PD-ResNet structure based on the ResNet unit was introduced in order to achieve the automated recognition of PD severity based on the H&Y scale (early PD: H&Y score $\leq$ 2.5, moderate to advanced PD: H&Y score > 2.5 [17]. In this study, polynomial enhanced dimensions technology was used to increase the dimensionality of the input data features, and synthetic minority over-sampling technique (SMOTE) was employed to accomplish sample balance. Further, PD-ResNet which gathers abundant feature information and effectively solves the gradient problem along with an enhanced focus loss function were suggested. Experiments demonstrated that the proposed PD-ResNet with enhanced focused loss function could identify H&Y stage efficiently with an accuracy of 92.00%.

Furthermore, researchers in [77] performed pairwise analysis of gait data and introduced a novel method for assessing the relative severity level of PD patients based on the scores of UPDRS scale (normal, mild, and moderate severity level). In order to achieve this objective, a novel deep learning architecture for pairwise rating of multivariate time-series data acquired by Ground Reaction Force (GRF) sensors worn on the foot was developed. In 10-fold cross validation, the proposed model, called Ranking by Siamese Recurrent Network with Attention, achieved 81.00% with an AUC of 0.88. To realize automated quantitative assessment of gait motor disorder in PD patients using gait videos, authors in [55]

proposed a two-stream spatial-temporal attention graph convolutional network (2s-ST-AGCN) under deep supervision and model-driven scheme. In particular, the spatial organization and temporal dynamics of the joints and bones were modeled. Experiments done on a clinical dataset to demonstrate the usefulness of the proposed model for identifying PD severity levels (MDS-UPDRS scores 0-4) yielded satisfactory results with an accuracy of 98.90%.

The review also identified two surveys utilizing advanced DL methods for PD recognition. By fusing and aggregating data from several sensors on the lower limbs, a unique hybrid model was suggested by [78] to discover the gait differences amongst three neurodegenerative disorders (ALS, PD, HD). Utilizing a spatial feature extractor (SFE), representative features of pictures or signals were generated. A novel correlative memory neural network (CorrMNN) architecture was developed to extract temporal characteristics from the two modalities' input in order to collect temporal information. The researchers then incorporated a multiswitch discriminator to link the observations with individual state predictions, achieving an accuracy of 99.00%. Finally, authors in [79] created a deep time series-based method for the detection of aberrant walking patterns in the gait dynamics of elderly individuals based on a hybrid LSTM-MLP network. The results demonstrated a testing accuracy of 70.00%.

2) Analysis of Upper Limb Motion: As already mentioned in previous Sections, motor symptoms occurring at upper limbs (fingers, wrists, hands, upper arms) are common manifestations of PD. Table IX in the Appendix summarizes the studies and their findings that employ DL methods for the detection of PD using motion characteristics presented in the upper limbs.

Looking at Section II, a typical manifestation of PD is the tremor symptom occurring at fingers, wrists and hands [80]. This type of tremor, also known as pathological hand tremor (PHT), compromises manual aiming, motor coordination, and movement dynamics and it is also considered as the main characteristic of the condition of essential tremor (ET) [81]. Effective treatment and management of the symptoms depends on the accurate and timely identification of the afflicted individuals, with the PHT features serving as a crucial parameter for differential diagnosis [33]. To this goal, the presented review identified two surveys that utilized advanced deep learning for discriminating between ET and Parkinson's tremor (PT). The respective studies are presented in Table V. Researchers in [82], integrated Gated Recurrent Unit (GRU) and LSTM algorithms. Initially, accelerometer sensors were utilized to capture hand tremors in each subject's three axial dimensions. These data are then pre-processed via the conventional scalar function and scaled in-unit variance, passed through the GRU model, and then served as input into the LSTM model to enhance its performance. Finally, they utilized a blockchain network to confirm the testing accuracy of the trained model, which was 74.10%. Additionally, in [83] a data-driven NeurD-Net model was proposed analyzing the kinematics of the hand classifying between PD and ET. NeurDNet was trained on more than 90 hours of hand motion signals including 250 tremor evaluations from 81 patients, exceeding its

state-of-the-art equivalents with a differential diagnostic accuracy of 95.55%.

Considering the detection of PD, researchers in [84] studied the applicability of CNN and CNN-BLSTM models using time series classification. Raw time series from pen-based signals were employed from the CNN-BLSTM and various data augmentation methodologies were presented in order to train these algorithms for PD detection on large-scale data. In this context, the Multi-Modal Collection (PDMultiMC) collection was produced [85], which contains recordings of online handwriting, voice signals, and eye movements. The HandPDMultiMC dataset, a subset of PDMultiMC, contains examples of handwriting from 42 participants (21 PD and 21 controls). Experimental results on the dataset revealed that CNN-BLSTM models trained with jittering and synthetic data augmentation provided the highest performance for early PD identification (97.62% accuracy) when combined with synthetic data augmentation.

3) Analysis of Speech: As a fundamental biological characteristic of humans, voiceprint is widely utilized in medical research and diagnostics, particularly in the identification of Parkinson's disease [89]. Despite the fact that there are several symptoms and features that signal PD, voice characteristics play a significant part among the predictive factors, as explained in Section II. A person with PD exhibits a variety of vocal impairments, including trembling and poor speaking. Voice analysis has the added advantages of being non-invasive, inexpensive, and easy to diagnose. The presented literature review identified a number of ardent researchers that invented new models and refined current ones to classify between PD and healthy subjects. The selected studies can been seen in Table V.

To address the restricted amount of current patient voiceprint datasets and samples, [86] developed a Spectrogram Deep Convolutional Generative Adversarial Network (S-DCGAN) for sample augmentation. S-DCGAN created a high-resolution spectrogram by adding network layers, the Spectral Normalization (SN) approach, and a feature matching strategy. To enhance the samples, spectrograms with high similarity and low distortion were chosen based on the Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR). Fréchet Inception Distance (FID) and GAN-train results demonstrated the data's potential to generalize. Moreover, the authors built the ResNet50 model with a Global Average Pooling (GAP) layer to efficiently collect and categorize voiceprint data in order to increase recognition accuracy. Finally, comparison tests on various models and classification techniques were performed with the results indicating that the S-DCGAN-ResNet50 hybrid model achieved the maximum voiceprint recognition accuracy of 91.25% and sensitivity of 92.5%, allowing it to discriminate between PD patients and healthy individuals more precisely than the simpler DCGAN-ResNet50.

Finally, in a study in [87], Shimmer, Jitter, Harmonic parameters, Frequency parameters, Detrended Fluctuation Analysis (DFA), Recurrence Period Density Entropy (RPDE), and Pitch Period Entropy (PPE) related features were employed and Conv-XGB had been selected to differentiate PD from

TABLE IV
SUMMARY OF SELECTED STUDIES EMPLOYING ADVANCED DL MODELS FOR PD-RELATED DIAGNOSTIC
PURPOSES BASED ON THE MODALITY OF GAIT

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Performance metrics	Modality
[70]	45 PD	TUG	FoG / not FoG	Automatic region localizaiton (SAN), end-to-end discrimination feature extraction, 5-CV	CNN + Attention	Acc. = .79, Sens. = .68, Spec. = .81	Gait
[71]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	Filtering, segmentation, SMOTE, HO	CNN-LSTM	Acc. = .94, Sens. = .96, Spec. = .95	Gait
[72]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	Multi-level discrete wavelet transform, segmentation	CNN-BiLSTM	Acc. = .90	Gait
[73]	21 PD (ON-OFF)	ADL	FoG / not FoG	Filtering, segmentation, temporal, frequency, MFCC feature extraction, LOSO-CV	CNN-LSTM-DNN	Sens. = .87, Spec = .87, AUC = .94	Gait
[74]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	Filtering, segmentation,10-CV	SE-CNN	Acc. = .96	Gait
[75]	45 PD	TUG	FoG / not FoG	Anatomic joint graph sequence, adjacency matrix estimation, Res-Net50 vertex, C3D vertex and C3D context feature extraction, 5-CV	GS-RNN	Acc. = .82, Sens. = 84, Spec. = .82, AUC = .89	Gait
[76]	7 PD	WALN	FoG / not FoG	Filtering, segmentation, transfer learning, data augmentation, LOSO-CV, 70/30	hybrid CNN	Sens. = .87	Gait
[77]	120 (PD + ALS)	WAL, SIT, STD	FoG / not FoG	Wireless Wi-Fi signals and channel state information extraction, radar technology and micro-Doppler signatures, 10-CV, HO	Autoencoder NN	Acc. = .98	Gait
[17]	296 PD (ON) / 161 HC	TUG, RWAL, WALN	PD severity (H&Y)	Polynomial elevated dimensions technology, SMOTE, improved focal loss, 10-CV, 90/10	PD-ResNet	Acc. = .92, Pr. = .94, Sens. = .90, Spec. = .94, f1 = .92	Gait
[78]	93 PD / 73 HC	WAL	PD severity (UP-DRS)	RSRNA architecture for pairwise ranking, attention learning mechanism, 10-CV	Siamese Recurrent Network with Attention	Acc. = .82, AUC = .89	Gait
[55]	142 PD	RWAL	PD severity (UP-DRS)	multi-scale spatial-temporal attention-aware mechanism spatial-temporal feature extraction, deep supervision strategy, 5-CV	2s-ST-GCN	Acc. = .99	Gait
[79]	25 ALS / 20 HD / 15 PD / 16 HC	RWAL	HC / PD / ALD / HD	Spatial feature extractor (SFE)	CorrMNN	Acc. = .99	Gait
[80]	60 (PD + HC)	WAL	PD / non-PD	Segmentation, temporal feature extraction	hybrid LSTM-MLP	Acc. = .70, Pr. = .65, Sens. = .70	Gait

WAL: Walking, RWAL: Roundtrip Walking, ADL: Activities of Daily Living, TUG: Times-Up-and-Go test, SIT: Sitting on chair, STD: Standing, WALN: Walking Narrow path, SMOTE: Synthetic Minority Oversampling TEchnique, HO: Hyperparameter Optimization, CV: Cross Validation, SAN: Spatial Attention Network, LOSO: Leave-one-subject-out, MFCC: Mel Frequency Cepstral Coefficients, ResNet: Residual network, C3D: 3D convolution, CorrMNN: Correlative memory neural network

healthy controls. Based on the results of this study, the Conv-XGB model provided better results than individually using CNN and XGBoost. Additionally, a stacked auto-encoder deep neural network architecture was presented to categorize the voice signals of Parkinson patients and healthy individuals [88]. As input to the stacked autoencoder deep network, a spectrogram and scalogram of voice signals were utilized. Support vector machine (SVM) and Softmax classifiers were utilized to evaluate the retrieved features and the stacked auto-encoder based time-frequency features with softmax classifier achieved the highest level of accuracy (87.00%).

4) Analysis of Facial Expressions: Expressions on the face represent the prevalent symptoms of Parkinson's disease. The majority of the time, medical professionals identify Parkinson's disease in patients through intrusive, costly, and arduous medical testing as well as careful overtime monitoring. Thus, it is vital to design an alternative, cost-effective, and lasting approach that can aid the physician in analyzing the entire behavior of PD patients [90]. However, based on the results of this review, only a few studies have utilized the modality of facial expressions to diagnose PD-related characteristics by integrating deep learning algorithms, as can be observed in Table XI in the Appendix. These algorithms follow more

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Performance metrics	Modality
[83]	17 PD / 15 ET	Hand poses (rest)	PT / ET	Standardization, blockchain network for model validation	GRU + LSTM	Acc. = .74	Upper limbs
[84]	47 PD / 34 ET	Hand poses	PT / ET	FFT, Hamming window, 75/25, 5-CV, HO (grid search)	NeurDNet (CNN+QDA)	Acc. = .96	Upper limbs
[85]	21 PD / 21 HC	Writing tasks	PD / HC	Jittering and synthetic data augmentation	CNN-BLSTM	Acc. = .98	Upper limbs
[87]	20 PD / 20 HC	Phonation exercises (/a/, /o/ and /u/)	PD / non-PD	Voice-related feature extraction, spectrogram computation, structural similarity index values and Peak Signal to Noise Ratio (PSNR) for data augmentation, global average pooling	S-DCGAN-ResNet50	Acc. = .91, Sens. = .92, Spec. = .90	Speech
[88]	23 PD / 8 HC	Phonation exercises	PD / non-PD	Shimmer, jitter, frequency and harmonic parameter feature extraction, DFA, RPDE, PPE, 3-CV	Conv XGB	Acc. = .88	Speech
[89]	50 PD / 50 HC	Phonation exercses, monologue, text reading, sentence expression, rapid repetition (/pa/, /ta/, /ka/)	PD / non-PD	Time-frequency representations, continuous wavelet transformation, auto-encoder-based feature extraction	SADN	Acc. = .87	Speech

HO: Hyperparameter Optimization, CV: Cross Validation, DFA: Detrended Fluctuation Analysis, RPDE: Recurrence Period Density Entropy, PPE: Pitch Period Entropy, S-DCGAN: Spectrogram Deep Convolutional Generative Adversarial Network, ResNet: Residual network

traditional DL architectures as typical CNNs dominate over other models such as LSTMs regarding the differential diagnosis of PD, the symptom of hypomimia and the emotional expressiveness.

B. DL for Non-Conventional PD-Related Diagnostic Targets

In this Section we focus on studies that do not concern common PD-related diagnostic targets as identified in Section II as well as our conducted literature review.

1) Vascular Parkinsonism: Vascular Parkinsonism (VaP) is often characterized by lower body parkinsonism with early and quick impairment of gait and/or postural control and less tremor, in contrast to Idiopathic Parkinsonism (IPD). Patients with VaP often present a characteristic shuffling gait, but may also exhibit significant FoG, even in early stages of the disease [91]. Two studies evaluated the effectiveness of some machine learning strategies in distinguishing IPD and VaP gait and are presented in Tables VII - I. In [92], two supervised machine learning techniques, Multiple Layer Perceptrons (MLPs) and Deep Belief Networks (DBNs), were utilized to undertake a comparative classification study. The decisional space consisted of gait characteristics, with or without neuropsychological test (Montreal cognitive assessment (MoCA) score), that were rated highest in an error incremental analysis. For the classification task of identifying parkinsonian gait by discriminating between patients (IPD+VaP) and healthy controls, both algorithms achieved excellent accuracy (93.00% with or without MoCA). In the

classification test of the two patient groups (VaP and IPD), the DBN classifier performed better (73.00% with MoCA) than the MoCA classifier. In 2021, a new approach for gait pattern differentiation that used CNNs based on gait time series with and without the influence of levodopa medication was proposed [93]. The gait data of VaP patients, IPD patients, and healthy people were collected using sensors worn on both feet. Utilizing the linear support vector machine, lasso, and random forest, recursive feature elimination was used to determine the ideal feature subset that led to the best outcomes. Multiple hyperparameters and feature subsets were utilized to implement CNNs. The best CNN classifier obtained an accuracy of 86.00% when the impact of levodopa medication (OFF/ON state) was considered simultaneously. The drug reaction to levodopa increased classification performance.

2) Multiple FoG Events: The combination of freezing of gait prediction and rhythmic laser signals may assist PD patients in overcoming FoG episodes. Two research sought to use the poor gait patterns, before FoG, to develop DL-based models for multi-class FoG prediction [94], [95]. For the accurate classification of the gait prior to FoG (pre-FoG), the slope of the impaired gait pattern was used to define the individualized pre-FoG phase. On the basis of the pre-FoG phase and the relabeled gait data, the customized labeled FoG prediction LSTM and CNN models were constructed, yielding positive performance results. In another research [96], using machine learning methods, the freezing of gait event was identified prior to its commencement, hence producing walking, FoG, and gait transition classes. Using the Boruta technique for

feature selection, the DMLP-based model was applied to 5-second windows with an average accuracy of 78.00%.

3) Speech Dysarthria Recognition: In the earliest stages of PD, 90% of patients experience voice abnormalities, specifically hypokinetic dysarthria [97]. Researchers in [39] examined the problem of speech-based categorization of amyotrophic lateral sclerosis (ALS) patients, Parkinson's disease patients, and healthy controls (HC). To this goal, a spectrogram-based method is utilized with a 2D-CNN. By feeding overlapping windows to the CNN, temporal elements were taken into account through the use of short signal segments or broad analysis filters. A categorization of dysarthria into three classes (ALS, PD, or HC) was conducted. In addition, authors conducted a classification experiment for PD severity (3 class). Both baseline Mel frequency cepstral coefficients (MFCC) data and log Mel spectrograms were utilized in experiments. For a variety of audio durations, classification results indicate that models trained on log Mel spectrograms regularly outperform MFCC's, achieving an accuracy of 93.00%.

Additionally, a multitask learning scheme was developed in [41] to evaluate the severity of several speech deficits in PD patients. Consideration was given to a CNN-based deep learning strategy for multitask learning. Time-frequency representations of segment transitions between voiced and unvoiced segments served as input to CNNs. The evaluated tasks corresponded to subscores of a comprehensive scale developed to assess the patients' dysarthria impairments. Multitask learning enhanced the generalization of CNN, resulting in more representative feature maps for assessing the speech symptoms of PD patients. The results suggested that training a CNN in a multitasks learning scheme is preferable to training individual CNNs to learn tasks for each deficiency of PD patients. The aforementioned research are presented in Table X in the Appendix.

4) Alterations in Facial Expressions - Hypomimia: In clinical practice, the evaluation of hypomimia symptoms remains subjective or is confined to the identification of a few landmarks that inadequately explain the disease's subtle manifestations [59]. A recent study [98] presented a novel digital biomarker, represented as a spatio-temporal convolutional representation that learns facial movement patterns to differentiate between Parkinson and control patients as seen in Table XI in the Appendix. The suggested architecture constructs a representation using 3D convolutional layers that are combined from inception modules, achieving salient face expression activation maps. In a retrospective investigation including 16 Parkinson patients and 16 controls, this method was verified. Using 480 video sequences, the architecture achieved an average accuracy of 91.87% in a disease classification condition task.

Further, researchers in [99], tested whether contemporary computer vision techniques may be used to detect veiled facial features and measure medication states in PD. In order to identify PD hypomimia signals, a CNN model was trained, using pictures collected from videos of PD patients and controls. This trained model was applied to clinical interviews with 35 PD patients in their drug-induced

and non-induced motor states. Conclusively, the algorithm detected PD hypomimia with a test set AUC of 0.71, compared to 0.75 for expert neurologists using the UPDRS-III Facial Expression score. In addition, the classification accuracy of the model for on and off drug states in clinical samples was 63.00% compared to 46.00% when utilizing clinical rater scores.

C. Fusion of Multi-Modal Information

Although preclinical Parkinson's disease detection has been investigated, a practical, cost-effective, and comprehensive screening diagnostic has not yet been developed [100]. Due to the high heterogeneity and complexity in the progression of PD, as well as the difficulties in collecting a single time-point measurement of a single sign, it would be almost impossible to fulfill the aim of precise treatment and severity evaluation without incorporating a combination of bio-signals from different modalities. In this work, we identified a number of studies that followed a multi-modal approach towards PD diagnosis and are depicted in Table VI.

Researchers in [63] proposed a unique classification method for PD patients and healthy controls utilizing Bidirectional Long Short-Term Neural Networks (BLSTM). SensHand and SensFoot inertial wearable sensors for upper and lower limb motion analysis were utilized to capture motion data for thirteen tasks drawn from the MDS-UPDRS Part III. The retrieved spatiotemporal and frequency characteristics were used to each participant as a single input for the development of a recurrent BLSTM to distinguish between the two groups. Maximum achievable accuracy was 82.40% and the respective findings demonstrated that the selected features greatly contributed to assessing the long-term pattern in BLSTM for the evaluation of the PD, and that the increase in batch size could affect the accuracy of the training and testing models. Furthermore, a DMLP classifier for mobile phone-based behavior analysis was presented in [105] to evaluate the course of PD patients by assessing their speech and movement patterns, as monitored by a smartphone accelerometer in their pockets at various times of the day. Popular machine learning classification algorithms were applied to a dataset from UCI and a dataset collected by the authors in order to classify each patient as Parkinson's positive or negative. In addition, the performance of each approach based on its ability to appropriately categorize patients into one of these groups was evaluated, thus proving that, DMLP outperformed the rest of the models on both datasets.

Research in [102] offered a multimodal investigation of the motor skills of PD patients, accounting for deep learning architectures based on time-frequency representations and CNNs that integrate information from speech, handwriting, and gait signals. The proposed method replicated the incapacity of patients to initiate and terminate movement of their lower and upper limbs and vocal muscles while carrying out the studies outlined in Table VI. The feature maps learned by the CNN trained with multimodal input allowed for the interpretation of the neural network's hidden representations. The initial convolutional layers of the CNN trained using time-frequency representations of speech revealed statistically significant differences between PD patients and controls. With

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Perfromance metrics	Modal.
[64]	64 PD / 50 HC	MDS-UPDRS tasks	PD / non-PD	Biomechanical feature extraction, normalization, standardization, zero-padding, 70/30, HO	recurrent BLSTM	Acc. = .82, Sens. = .92, Spec. = .76	gait, upper limbs
[102]	20 PD/ 20 HC	ADL, WAL, SIT, STD	PD / non-PD	Filtering, segmentation, normalization, standardization, 70/30, shimmer, jitter, spread, acceleration time series feature extraction, leave-one-out 10-CV	DMLP	Pr. = .98	Speech, gait
[103]	44 PD(ON) / 40 HC	DDK, Writing and drawing, WAL	PD severity	Transitions identification with FFT, extended Geneva minimalistic acoustic parameter set, hand kinematic features, gait kinematic feature extraction, 80/10/10, HO (Bayesian tuning), Fusion: Emdeddings from the 3 bio-signals at the last hidden layer of each CNN are concatenated towards multi-modal vector formation serving as input to a radial basis SVM	CNN	Acc. = .56	Speech, gait, upper limbs
[104]	87 PD / 46 HC	AFT, WAL, Voice test, memory test	PD / non-PD	Segmentation, rhythm, spatial variability, and fatigue, temporal and frequency, total score, number of levels attempted, and number of incorrect responses feature extraction, 10-CV, Individual Source Model Evaluation, Classifier Ensemble	RF + LR + DNN + CNN	Acc. = .82	Speech, gait, upper limbs
[105]	88 PD / 46 HC	AFT, WAL, Voice test, memory test	PD / non-PD	Filtering, change point detection algorithm, TCN feature extraction, Temporal Encoder with Multi-modal Attention, Embedding Self-Attention Pooling, 5-CV	RNN	AUC = .79, f1-score = .82	Speech, gait, upper limbs

WAL: Walking, ADL: Activities of Daily Living, SIT: Sitting on chair, STD: Standing, HO: Hyperparameter Optimization, CV: Cross Validation, AFT: Alternate Finger Tapping, DDK: diadochokinetic analysis, TCN: Temporal Convolutional Networks

the final handwriting-trained CNN layer, comparable results were obtained. The combination of the three bio-signals was the most reliable method for classifying PD patients according to their disease stage. CNNs seemed ideal for modeling the problems of Parkinson's disease (PD) patients to initiate and terminate the movement of separate limbs, allowing for the correct categorization of PD patients and control participants.

Combined dataset deconstruction with multi-source ensemble learning enabled participants with incomplete data to be included in the training of machine learning models, as demonstrated in [103]. Using multi-source ensemble learning in conjunction with CNNs that capitalize on the quantity of the available data, researchers achieved an accuracy of 82.00% in PD classification, 9.00% more compared to conventional strategies. The rise in accuracy was attributed in part to the use of CNNs with a DNN and in part to the development of models employing a large cohort of participants. Finally, [104] presented a unique time-series based on deep learning method to Parkinson's Disease prediction using remotely and irregularly acquired speech, hand motion, and gait data from smartphones. Using the Neural Ordinary Differential Equations, researchers synced discrete data to unified observational time points in order to generate multimodal time-series representations. In both the temporal and modality aspects, two hypothesized attention processes acquired key characteristics from noisy signals in an adaptive manner. The success of the suggested method was shown by insights and better quantitative and qualitative outcomes on a large public dataset.

D. Sensory Equipment for Detection of PD-Related Clinical Signs

1) Sensors Capturing PD Gait: The analysis of gait, and specifically the analysis of the various events taking place during the stance phase and swing phase, is essential for the treatment of a variety of orthopaedic and neurological disorders [106]. The detection of typical gait events is a helpful technique for improving gait analysis, evaluating therapies for abnormal gait, and developing devices and sensors for gait support [22], [107], [108]. Therefore, gait analysis requires the quantified study of the parameters related to force, time and distance, by calculating a range of important temporal and spatial characteristics [109].

To analyze key features and movement patterns in normal and pathological gait, a variety of methods for computer-assisted analysis have been developed and are in clinical use. Human gait characteristics are detected by fixed devices such as optical motion capture systems, floor-based force platforms, and electronic treadmills, which are considered as "golden standards" [110]. However, these systems are only suitable for hospitals or hospital facilities due to their size, high cost, and the need for specialized personnel.

Video-based motion capture systems and instrumented motion analysis systems, have been well studied for obtaining gait characteristics. From the total forty-four (46) studies, 20% utilized COTS camera equipment in order to capture the body poses and motion of the subjects providing the researchers with precise lower limb markers for gait analysis. However the

value of the camera-based setups has been mostly limited to a laboratory context, they are expensive and require specialized movement laboratories [111].

In alleviating the limitations of camera-based systems, wearable motion sensing devices and systems have been developed for accurate gait analysis in daily life settings. In particular, recent technological advancements in microelectromechanical systems (MEMS) have enabled the design and development of lightweight and low-cost inertial measurement units (IMUs) based wearable sensors that are used for gait and balance monitoring [112], [113]. The authors in [114] report results that indicate the superiority of IMUs over camera-based setups regarding reliability and precision for the detection of PD based on gait related features. Therefore, it is well-established that mobile inertial measuring equipment, i.e. accelerometers, gyroscopes and magnetometers, are capable of objectively tracking gait motion [115], [116]. Based on the results of our literature review, almost half of the selected surveys (45%) capitalized on the accuracy and reliability offered by IMUs and consequently used them to analyse gait and extract gait related characteristics. More specifically, the aforementioned studies used data obtained from IMUs placed primarily on the trunk (waist, lower back, chest), ankles and thighs at rates of 33, 31 and 20% respectively, according to the present literature review. Other sites where IMUs were placed were the wrists, calves and feet. Fig. 4 presents the placement and the percentage of the studies that employ IMU sensors on different lower extremities sites.

Nevertheless, due to the well-known defects associated with inertial sensor measurement data, such as time-variant sensor biases and measurement noise [117], there is an increasing need to exploit other types of sensors integrated in wearable devices. One such type of sensors are insole-based plantar pressure monitoring systems, that have risen to prominence as a key wearable tool for monitoring the course of chronic disease in patients. Internal plantar pressure measurement insoles and foot switches are the cutting-edge technology for gait phase analysis, since each phase can be associated with a specific value of the sensor output. Insoles offer better resolution compared to foot switches, as they allow the recording of the whole foot contact with the ground, which is not affected by the positioning of the foot switch [107], [118]. In addition, an important advantage is the ease of use as minimal intrusion is achieved.

Consequently, insole-based plantar pressure monitoring systems are expanding rapidly around the globe, with several research institutes and businesses demonstrating an increased interest in the field [119]. In gait analysis, they are frequently used to count steps and extract spatiotemporal information. They are regularly used in stability investigations to identify the center of pressure and, consequently, postural stability [120]. Although plantar pressure insoles have proven to be reliable for the analysis of gait and walking patterns, only 30% of the total number of selected studies have utilized insole systems, possibly due to their higher cost concerning hardware equipment but also software licences. Interestingly, none of the studies combined two or three of the aforementioned technologies (cameras, IMUs, insole pressure sensors). Finally,

Fig. 4. The location of sensors for the assessment of PD severity and tremor based on gait.

in 5% of the selected studies the sensory equipment was not specified.

Typically, while normal gait steps have a main frequency from 0.5 to 3 Hz, FoG events present a main frequency to the range of 6 to 8 Hz [121]. Therefore, it is essential to note, that the maximum and minimum sampling frequencies between all IMUs and pressure sensors used in the selected studies were found to be between 1080 and 30 Hz respectively, whereas the maximum and minimum frames per second between all the studies that employed camera-based equipment were 240 and 25 fps respectively.

2) Sensors Capturing Upper Limp PD Motion and Tremor: Considering the importance of studying the motion of upper limbs towards the detection of PD and its evolution, the majority of the studies selected in this review utilized IMUs for PD detection, estimation of PD severity or the identification of Parkinson's tremor (PT). Specifically, in 44% of the studies the IMUs were placed on the subjects' wrists, in 25% they were placed on the subjects' fingers and the metacarpal area, and in 6% of the studies they were placed on the area, as shown in Fig.5. The rest of the studies did not clarify the exact point of placement of the IMUs. Furthermore, in addition to using IMUs, the hand position during clinical evaluation tasks was captured from a 3-camera setup in [122].

Fig. 5. The location of sensors for the assessment of PD severity and tremor based on the movement of the upper limbs.
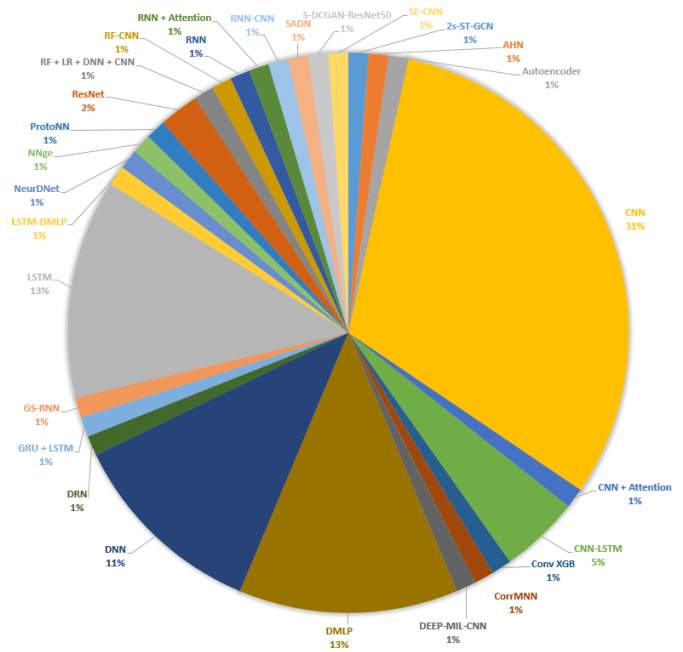

Fig. 6. Pie chart representation of various deep learning models used in PD diagnosis based on gait, upper limb, speech and facial expressions modalities.

Additionally, a number of studies used digitized tablets to locate the coordinates of the pen during handwriting tasks and smartphones placed on a desk to analyze finger tapping activity.

Concluding, the characteristics of Parkinson's disease tremors have been thoroughly investigated. It has been reported that the frequencies of the classical rest tremor, isolated postural tremor, and kinetic tremor during slow movement are 3-7 Hz, 4-9 Hz, and 7-12 Hz, respectively [123]. The sensory equipment used in the relevant reviewed studies obtained motion information from the upper limb extremities with sampling rates ranging from 40 Hz to 12 kHz, enough to satisfy the Nyquist theorem.

3) Sensors Capturing PD Speech: As speech is considered of major importance for the early detection of PD patients, the way of recording the subjects' voice is a crucial issue in addition to the appropriate selection of the tasks to be performed. As PD patients suffer from a variety of neurological symptoms, the recording should be more convenient. As a consequence, initially, the data are usually recorded at 44.1 kHz or 48 kHz [124], [125]. The choice of the recording devices includes cell phones, so as to make the recording procedure available at any environment [39]. Moreover, in other studies, systems consisting of extend USB sound card with low level noise combined with microphones of "clip-on" type are used [42]. Furthermore, notice that, the understanding of neurological disorders through speech requires not only recordings of good quality but also of a satisfying duration for their accurate analysis, usually not very long. Finally, for the creation of a well-defined dataset, important role play the

age and the balanced number of male and female subjects' choice [40], [41].

4) Sensors Capturing PD Facial Expressions: Expressions on the face represent the prevalent symptoms of Parkinson's disease. The majority of the time, medical professionals identify Parkinson's disease in patients through intrusive, costly, and arduous medical testing as well as careful overtime monitoring. Thus, it is vital to design an alternative, cost-effective, and lasting approach that can aid the physician in analyzing the entire behavior of PD patients [90]. In the studies reviewed, researchers record the facial expressions of participants by using high resolution digital cameras in order to identify relevant biomarkers and patterns that link the hypomimia with PD.

VI. DISCUSSION

The primary objective of this survey is to examine and project future research directions in the field of DL-based PD diagnosis techniques. This review provides a comprehensive overview of advanced deep learning based approaches for PD manifestations prediction by utilizing data from different modalities and in particular gait, speech, upper limbs and facial movements. We have analysed the importance of collecting physiological signals from these different modalities to diagnose Parkinson's Disease.

Based on the results presented in this literature review, almost one third (30%) of the total investigated studies made use of CNNs to reach their diagnostic targets, as depicted in Fig. 6. Moreover, CNNs were additionally employed in combination with other DL algorithms such as LSTMs or attention models. LSTM networks were the second most widely used models while a vast majority of the studies made use of simpler neural network models such as deep multi layer
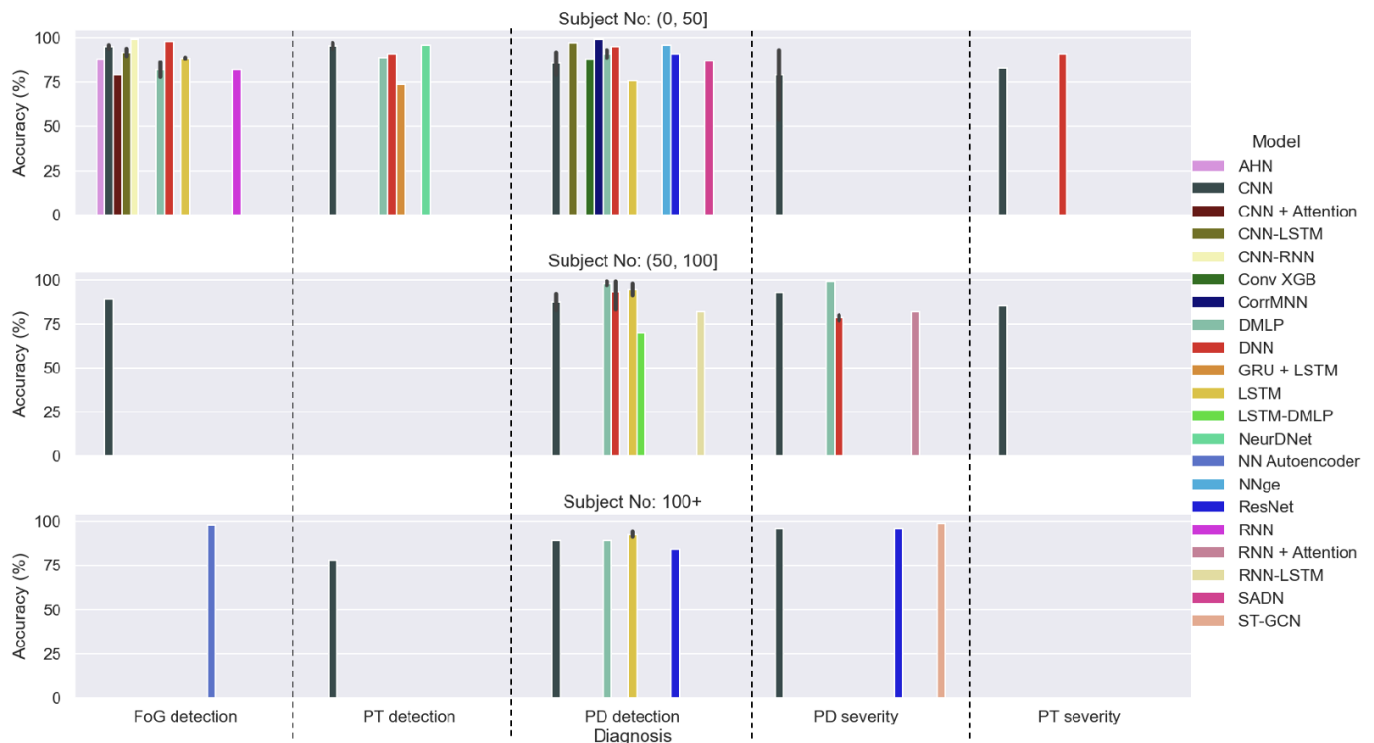
Fig. 7. Diagrams representing the accuracy rates of utilized DL methods. The thin black line refers to the standard deviation of the accuracy, in the cases where same models were used for the same classification target.

perceptrons. Additionally, 11% of the selected studies followed the more classical deep neural network concept to achieve their diagnostic targets. In addition, most articles considered accuracy, specificity, and sensitivity as metrics to validate the developed PD classification models' performance.

Observing Fig. 7, three (3) diagrams can be distinguished which vary according to the number of participants in each study. Each graph shows the accuracy rate of the applied model in relation to the 5 main classification targets as identified in the literature review i.e. FoG detection, PT detection, PD detection, PD severity and PT severity. The highest success rates were observed in studies that carried out analysis of gait and reached accuracy rates of over 98.00%, thus highlighting the efficiency of this particular modality in PD identification. These studies focused on the identification of PD and FoG events, but also on the discrimination of the level of severity of the disease. In fact, two of these studies involved more than 100 volunteers, which may have intrigued researchers to explore and investigate the potential of more complex DL models such as Autoencoders and two-stream spatial-temporal attention graph convolutional network (2s-ST-AGCN). Interestingly, among the studies with the highest accuracy scores and despite the widespread use of CNNs in the selected studies, only one of them used this technique and even in combination with RNN, towards FoG detection.

Regarding the studies that dealt with the detection of disease characteristics using upper extremity movement data, and by looking at Fig. 7, we can observe that research on the upper extremities has focused either on the recognition of the disease or on the identification of PT and classification of its severity. The recognition of the disease was performed successfully at a rate of 97.00% by the hybrid CNN-LSTM model and

by a DMLP neural network. With the same percentage, the existence of the PT was also identified by the CNN algorithm. Interestingly, the classification of the severity levels of tremor was successfully performed at rates between 83.00 and 91.00%, which are among the lowest success rates of the upper limbs related studies investigated.

The majority of the research that used the speech modality aimed at identifying or differentially diagnose PD. According to the results of this literature review, the highest recognition rate was achieved by the NNge model and reached 96.00%. The classification of the disease severity levels was similarly high. More specifically, the levels of dysarthria were classified at a rate of 93.00% by the CNN model. Finally, novel network architectures such as ResNet, Conv XGB and SADN achieved a high rate in disease recognition above 87.00%.

The difficulty in altering facial expressions and the phenomenon of hypomimia are a result of the muscle stiffness caused by Parkinson's disease. Thus, the CNN model was able to identify among PD and non-PD patients at a rate of 92.00%, while in another study with the same diagnostic target the LSTM model achieved a 76.00% rate.

Without a combination of biosignals from various modalities, it would be nearly impossible to achieve the goal of precise treatment and severity evaluation, as previously stated. As mentioned, the highest accuracy rates in the single modality studies were related to the binary diagnosis of the disease. In accordance to this, the highest accuracy rate in the selected multi-modality studies (82.00%) is found when attempting to classify between PD and non-PD subjects. Notably, the two models that achieved this rate were the recurrent BLSTM model and the DL architecture combining DNN and CNN models. In fact, we must highlight the importance of a large

TABLE VII
SUMMARY OF SELECTED ARTICLES FROM THE LITERATURE ON PD AND ITS SYMPTOMS DIAGNOSIS
BASED ON LOWER LIMB MOVEMENT AND GAIT

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Performance metrics
[127]	24 PD (OFF)	ADL, RWAL, WALN	FoG / not FoG	Filtering, segmentation, stepwise temporal and frequency feature extraction method, mutual information-based feature selection	DMLP	Acc. = .86
[128]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	Segmentation, temporal feature extraction, DT for ranking feature importance, SVM-RFE feature selection, 70/30	AHN	Acc., Sens., Pr. = .88
[129]	67 PD	TUG	FoG / not FoG	Morlet wavelet transform, segmentation, 80/20, model regularization	2D-CNN	Acc. = .89, Sens. = .82, Spec. = .96
[130]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	RNN and LSTM comparison, transfer learning, 50/50	LSTM	Acc. = .88
[56]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	Segmentation, comparison with CNN, LSTM, ResNet-50	DRN	Sens. = 86, Spec. = .94
[131]	38 PD (ON) / 21 HC	RWAL	FoG / not FoG	Filtering, segmentation, temporal and frequnecy feature extraction, Pearson feature selection, 10-CV	DNN	Acc. = .98
[132]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	Filtering, segmentation, 10-CV	CNN	Sens. = .93, Spec. = .87
[28]	4 PD		FoG / not FoG	Full-wave rectification, fitlering, segmentation, Daubechies wavelet transform, peak detection algorithm, time, spatiotemporal, physiological feature extraction, stratified 3-CV, HO (grid search)	CNN	Acc. = .94
[133]	118 PD / 21 HC	ADL	FoG / not FoG	Segmentation, multi-input network	CNN	Sens., Spec. = .88
[134]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG	Temporal and frequency feature extraction, 10-CV	ProtoNN	Sens. = .95, Spec. = .99
[135]	10 PD	RWAL	FoG / not FoG	Segmentation, 75/25	2D CNN-RNN	Acc. = .99, Sens. = 97, Spec. = .99, Pr. = .89
[136]	11 PD (OFF)	RWAL, WALN	FoG / not FoG	Segmentation, CoP and GRF feature extraction, one-freezer-held-out cross validation	LSTM	Sens. = .72, Spec. = .81
[137]	18 PD	RWALN	FoG / not FoG	Segmentation, SMOTE, ADAptive SYNthetic technique, statistical feature extraction, LOSO-CV, 60/40	DMLP	Sens. = .82, Spec. = .94, f1-score = .88 (p-d), Sens. = .68, Spec. = .90, f1-score = .64 (p-i)
[96]	5 PD (ON)	RWAL+, WALN+	FoG / not FoG / pre-FoG	Filtering, segmentation, basic signal and statistical feature extraction, ANOVA and mutua information future selection, leave-one-subject-out 10-CV, 70/30	CNN	Sens. = .92, .92, .94, Spec. = .98, .96, .95
[95]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG / pre-FoG	Filtering, Hampel outlier identifier, polynomial interpolation, segmentation, 80/20	LSTM	Acc. = .89, Sens. = .90, Spec. = .87
[97]	10 PD (OFF)	WAL, RWAL, ADL	FoG / not FoG / Transition	Segmentation, temporal and frequency feature extraction, Boruta algorithm for feature selection, 70/30	DMLP	Acc. = .78, Sens. = .72, Spec. = .85 (FoG detection)
[138]	93 PD / 73 HC	WAL	PD / non-PD	Normalization, statistical feature extraction, 70/20/10	LSTM	Acc. = .98, f1-score = .98
[139]	17 ALS / 20 HD / 15 PD / 16 HC	RWAL	PD / non-PD	Filtering, SMOTE, phase synchronization with Hilbert transform, PCA, temporal feature extraction	DMLP	Sens. = 1.00, Spec. = .81, AUC = .93

WAL: Walking, RWAL: Roundtrip Walking, ADL: Activities of Daily Living, TUG: Times-Up-and-Go test, WALN: Walking Narrow path, RWAL:
Roundtrip Walking + cognitive task, WALN+: Walking Narrow path + cognitive task, DT: Decision tree, PCA: Principal Component Analysis,
SVM: Suppor vector machine, RFE: Recursive feature elimination, ANOVA: Analysis of variance, SMOTE: Synthetic Minority Oversampling
TEchnique, HO: Hyperparameter Optimization, CV: Cross Validation, LOSO: Leave-one-subject-out, ResNet: Residual network
Studies presented in Table IV are not presented in this Table.

TABLE VIII

SUMMARY OF SELECTED ARTICLES FROM THE LITERATURE ON PD AND ITS SYMPTOMS DIAGNOSIS BASED ON LOWER LIMB MOVEMENT AND GAIT (CONTINUE)

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Performance metrics
[140]	40 PD (OFF) / 40 HC	WAL	PD / non-PD	Kalman filter, statistical feature extraction, segmentation, 10-CV	DMLP	Acc. = .89, Sens. = .89, Spec. = .89, f1-score = .87
[141]	93 PD / 73 HC	WAL	PD / non-PD	Phase space reconstruction, Empirical mode decomposition, GRF positions feature extraction, LOSO-CV	DNN	Acc. = .99
[142]	93 PD / 73 HC	WAL	PD / non-PD	n/a	Stacked CNN	Acc. = .89
[143]	93 PD / 73 HC	WAL	PD / non-PD	Layer-wise relevance propagation, 60/20/20, HO	2D-CNN	Acc. = .95
[144]	88 PD / 94 HC	WAL	PD / non-PD	90/10, 5-CV, LSTM, 1D-CNN, CNN+LSTM comparison	1D-CNN	Acc. = .83, Pr. = .80, Sens. = .76, f1-score = .78
[145]	11 PD / 11 HC	RWAL	PD / non-PD	Spatiotemporal feature extraction, retro-propagation process fro saliency maps, tracking of maximum activation values in the previous dense layers and pseudo-deconvolution process for feature selection, leave-one-out CV	CNN	Acc. = .95
[146]	93 PD / 73 HC	WAL	PD / non-PD, PD severity	10-CV, HO	DMLP	Acc. = .99, Acc. = .99
[147]	76 PD (ON) / 67 HC	WAL	PD / non-PD, PD staging	Spatial, spatiotemporal, kinematic feature extraction, PCA, 80/10/10	DNN	Acc. = .98, Sens. = .89, Spec. = .89, Acc. = .77, Sens. = .77, Spec. = .91
[148]	93 PD / 73 HC	WAL	PD severity	Spatiotemporal feature extraction, segmentation, 18 parallel 1D-Convnets	CNN	Pr. = .87, Sens. = .85, f1-score = .85
[149]	18 PD / 42 HC	TUG	PD severity	Mask R-CNN for body area detection, spatial and temporal feature extraction, temporal and spatial fusion module, 3-CV	CNN	Acc. = .91, Pr. = .92, Sens. = .91
[94]	14 VaP / 15 iPD / 36 HC (OFF-ON)	RWAL	vap / iPD / non-PD	Multiple regression normalization, Spatial, temporal, foot clearance feature extraction, RFE throug Linear Kernel SVM and Lasso feature selection, 5-CV, HO	RF-CNN	Acc. = .79 (OFF), Acc. = .82 (ON), Acc. = .86 (ON and OFF)
[93]	15 VaP / 15 iPD / 15 HC	RWAL	vap / iPD / non-PD, VaP / IPD	Normalization, temporal, spatial, spatiotemporal feature extraction, Kruskal-Wallis test, Mann- Whitney test with Bonferroni correction feature selection, 75/25, 10-CV	DMLP, DBN	Acc. = .93, Acc. = .73
[150]	32 PD / 16 HC	RWAL	Advanced PD / not-adv. PD, early-PD / non-PD	5-CV	CNN	Acc. = .93, Sens. = 1.00, Spec. = .90, Acc. = 1.00, Sens. = .99, Spec. = 1.00
[151]	524 PD / 43 ET	iSAW	PD / ET	Kinematic, spatial, temporal, spatiotemporal feature extraction, SMOTE, stratified 3-CV	DMLP	Acc. = .89, Pr. = .61, Sens. = ..61, f1-score = .61
[152]	10 MS / 9 PD / 14 HC	WAL	PD / MS / HC	Segmentation, 3D joint keypoints extraction, 5-CV	CNN	Acc. = .79, AUC .93

WAL: Walking, RWAL: Roundtrip Walking, PCA: Principal Component Analysis, SMOTE: Synthetic Minority Oversampling TEchnique, HO: Hyperparameter Optimization, CV: Cross Validation, LOSO: Leave-one-subject-out, ResNet: Residual network
Studies presented in Table IV are not presented in this Table.

database, as the highest rate of classification of disease was found in two studies with over 100 participants, while the lowest (56.00%) was found in a survey with 84 subjects.

A. Challenges and Limitations

Despite the significant contributions made through a comprehensive synthesis of the most pertinent information on deep learning methods for clinical diagnosis, this systematic review has some limitations. It is acknowledged that collecting actual patient data is the most difficult task in the healthcare sector compared to other research fields. In most instances, neurodegenerative disease-related medical datasets are imbalanced. Specifically, deep learning studies for each investigated modality (gait, upper limb motion, speech, facial expressiveness) may utilize either public or private datasets to train their models. Therefore, comparing the performance of

TABLE IX
SUMMARY OF SELECTED ARTICLES FROM THE LITERATURE ON PD AND ITS SYMPTOMS DIAGNOSIS BASED ON UPPER LIMB MOVEMENT

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Performance metrics
[153]	21 PD / 21 HC	Hand poses (rest)	PD / HC	Wavelet paket decomposition, Energy, Mean value, Variance, Entrop feature extraction	DNN	Acc. = .95
[154]	PD / HC	Typing procedure	PD / HC	Stratified k-CV, HO (grid search)	LSTM	Acc. = .76
[58]	58 PD / 29 HC	Writing tasks	PD / HC	2 stacked LSTMs	LSTM	Acc. = .91, Sens. = 1.00, Spec. = .65, f1-score = .94
[85]	21 PD / 21 HC	Writing tasks	PD / HC	Jittering and synthetic data augmentation	CNN-BLSTM	Acc. = .98
[155]	55 PD / 45 YHC / 49 EHC	Writing tasks	PD / Young HC, PD / Eldery HC	Kinematic, non-linear dynamics and neuromotor features extraction, meta-parameter optimization, leave-ne-out CV, grid search	DMLP	Acc. = .97, Acc. = .78
[156]	43 PD	Hand poses (rest, posture)	PD severity	Filtering, Windowing, Frequency and amplitude related features extraction	DNN	Acc. = .91, Sens. = .90, Spec. = .91
[157]	398 (PD or ET)	Hand poses (resting, stretching, winging, and vertically winging)	PT / ET	Frequency and amplitude features extraction, HO (grid search), 80/20, 10-CV,	CNN	Acc. = .78, Spec. = .67, Sens. = .86, Pr. = .79, f1-score = .83, AUC = .77, FPR = .33
[83]	17 PD / 15 ET		PT / ET	Standardization, blockchain network for model validation	GRU + LSTM	Acc. = .74, Comm. Cost = 448 bits, Comp. cost = 0.056 ms
[84]	47 PD / 34 ET	Hand poses (rest, postural, weight-holding, finger-to-nose motion)	PT / ET	FFT, Hamming window, 75/25, 5-CV, HO (grid search)	Neur-DNet (CNN + QDA)	Acc. = .96
[158]	17 PD	Hand poses (rest, postural)	RT / PT	Daubechies4 wavelet function, windowing, time and frequency features extraction, tremor eigenvalues extraction, PCA	DNN	Acc. = .95 (RT) Acc. = .91 (PT)
[57]	5 PD / 5 HC	Walking tasks	PT / No PT	Filtering, statisitical features extraction	CNN	Acc. = .97, Sens. = .92, Spec. = .97, f1-score = .94, time = 80 10-6 s
[159]	30 PD	Clinical evaluation tasks	PT / No PT	Segmentation, 5-CV, leave-one-out evaluation	CNN	Acc. = .94, Pr. = .89, Sens. = 1.00 , f1-score = .94
[160]	6 PD	ADL, Clinical evaluation tasks	PT / No PT	CNN trained on tremor and activity spectra	DMLP	Acc. = .89
[161]	20 PD (ON) / 20 HC	Writing tasks	PT severity	HOG and 2D-CNN feature extraction, stratified 5-CV	1D CNN	Acc. = .83, Sens. = .85, Spec. = .81, AUC = .91
[162]	92 PD	Hand poses (rest)	PT severity	Filtering, FFT	CNN	Acc. = .85, Kappa = .85, Corr. = .93, RMSE. = .35
[163]	31 PD / 14 HC	ADL in the wild	PT severity	Filtering, linear interpolation, segmentation, Leave-one-out CV	DEEP-MIL-CNN	Pre. = .99, Sens. = .90, Spec. = .99, f1-score = .94
[164]	24 PD (OFF)	ADL	PT severity	Filtering, windowing, frequency related features extraction, Wilcoxon rank sum feature selection, subject-based leave-one-out CV, HO (grid search)	LSTM	r = .77, MAE = 1.32

HO: Hyperparameter Optimization, CV: Cross Validation, FFT: Fast Fourier Transform, PCA: Principal Component Analysis, HOG: Histogram of Oriented Gradients, PT: Parkinson's Tremor, ET: Essential Tremor, GRU: Gated Recurrent Unit
Studies presented in Table V are not presented in this Table.

two deep learning models that were not trained with the same dataset could be quite challenging.

Furthermore, due to the significant variability of the studies with respect to the case of data and presentation of results, it was difficult to directly compare the outcomes associated with each type of model across studies, as some studies failed to indicate whether model performance was evaluated using a test set, and/or results given by models that did not yield the best per-study performance. In addition, it is our view that there is a lack of research on facial movement and expressiveness, as well as analyses involving multiple modalities. This makes it difficult to identify the model that performs better in these two categories. Lastly, the vast number of deep learning models proposed for gait

TABLE X

SUMMARY OF SELECTED ARTICLES FROM THE LITERATURE ON PD AND ITS SYMPTOMS DIAGNOSIS BASED ON SPEECH

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Performance metrics
[88]	23 PD / 8 HC	Phonation exercises	PD / non-PD	Shimmer, jitter, frequency and harmonic parameter feature extraction, DFA, RPDE, PPE, 3-CV	Conv XGB	Acc. = .88
[42]	51 PD / 27 HC	Text reading and monologue	PD / non-PD	Speech rhythm and prosodic feature extraction, LOSO-CV	DNN	Acc. = .83, Sens. = .79, Spec. = .79 (monologue exercise)
[165]	23 PD / 8 HC	Phonation exercises	PD / non-PD	Shimmer, jitter and F0 related feature extraction, SMOTE, 10-CV, HO (random search)	NNge	Acc. = .96
[166]	23 PD / 8 HC	Phonation exercises	PD / non-PD	Shimmer, jitter and F0 related feature extraction, denoising, PCA feature reduction, HO, 10-CV	Opt-NN	Acc. = .95, Pr. = .95, Sens. = .95, f1-score = .95, MSE = .05, MAC = .05
[167]	188 PD / 64 HC	Voice recordings	PD / non-PD	TQWT, TF, MFCCs, WT, VF feature extraction. data normalization, feature mapping and convertion to RGB, 10-CV, 70/30	LSTM	Acc. = .94, Sens. = .96, Spec. = .96, Pr. = .91, f1-score = .93
[62]	564 PD / 192 HC & 23 PD / 8 HC	Phonation exercises	PD / non-PD	Oversampling technique, 10-CV, HO (random search), 90/10	LSTM	Acc. = .91
[87]	20 PD / 20 HC	Phonation exercises (/a/, /o/ and /u/)	PD / non-PD	Voice-related feature extraction, spectrogram computation, SSIM values and PSNR for data augmentation, Global Average Pooling	S-DCGAN-ResNet50	Acc. = .91, Sens. = .92, Spec. = .90
[40]	41 PD / 40 HC	Phonation exercises (sustained /a/)	PD / non-PD	Spectrogram analysis. data downsampling, windowing (overlapping), removal of silent speech parts, 75/15/15, 10-CV	DCNN	Acc. = .76
[39]	60 PD / 60 HC / 60 ALS	SPON, IMAG, PHON, DIDK	PD / HC / ALS, PD speech severity	Windowing, log MEL spectrograms computation, 5-CV, 80/20	CNN	Acc. = .93 (DIDK), Acc. = .87 (DIDK)
[89]	50 PD / 50 HC	Phonation exercses, monologue, text reading, sentence expression, rapid repetition (/pa/, /ta/, /ka/)	PD / non-PD	Time-frequency representations, SFTF-CWT, auto-encoder-based feature extraction	SADN	Acc. = .87
[38]	80 PD	Repetition of sustained vowels (/a/, /i/ and /u/), text reading, dialogue	PD speech severity	ASR feature extraction), eGeMAPS, Pearson feature selection	DNN	Acc. = .80
[41]	50 PD / 50 HC	Sustained phonation, word, and monologue, text reading, sentence expression, rapid repetition (/pa/, /ta/, /ka/)	Dysarthria severity	Time-frequency representations, multi-task learning	CNN	Acc. = .54
[168]	188 PD / 64 HC	Voice recordings	PD / non-PD	Time-frequency, MFCCs and Delta features in MFCC, baseline, energy and entropy based feature extraction, SMOTE	ResNet-50	Acc. = .84
[169]	23 PD / 8 HC	Phonation exercises	PD / non-PD	Data cleaning, feature scaling	DNN	Acc. = .95

SPON: Spontaneous speech/ monologue, IMAG: describing images, PHON: Sustained Phonation, DIDK: Diadochokinetic tasks, HO: Hyperparameter Optimization, CV: Cross Validation, LOSO: Leave-One-Subject-Out, SMOTE: Synthetic Minority Oversampling TEchnique, DFA: Detrended Fluctuation Analysis, RPDE: Recurrence Period Density Entropy, PPE: Pitch Period Entropy, PCA: Principal Component Analysis, WT: wavelet transform, TF: time-frequency, TQWT: tunable Q-factor wavelet transform, MFCCs: Mel frequency Cepstral coefficients, VF: vocal fold, SSIM: Structural Similarity Index, PSNR: Peak Signal to Noise Ratio, SFTF-CWT: Short time Fourier transformation - Continuous wavelet transformation, ASR: Automatic Speech Recognition, S-DCGAN: Spectrogram Deep Convolutional Generative Adversarial Network Studies presented in Table V are not presented in this Table.

analysis makes it difficult to identify the most effective model.

VII. CONCLUSION

Parkinson's disease requires early diagnosis and treatment to minimize its impact and preserve patients' independence. Early and accurate clinical diagnosis of PD is especially crucial in the context of emerging neuroprotective treatments. While prodromal diagnosis primarily relies on non-motor symptoms, our systematic review proposes a multi-modal deep learning approach to enhance early clinical diagnosis based on motor signs. By integrating data from various sources, DL-based solutions aim to improve the accuracy of clinical

TABLE XI
SUMMARY OF SELECTED ARTICLES FROM THE LITERATURE ON PD AND ITS SYMPTOMS DIAGNOSIS BASED ON FACIAL EXPRESSIONS

Ref.	Subjects	Experiment(s)	Diagnosis	Development process	Model	Performance metrics
[99]	16 PD / 16 HC	Phonation exercise	PD/ non-PD	Inclusion of inception-like 3D layers, embedding representation, LOSO-CV, HO	CNN	Acc. = .92, Pr. = .96, Sens. = .87, f1-score = .91, AUC = .95
[170]	70 PD	Expression mimicing	Normal / happy / angry / sad	Face detection, 80/20, HO	CNN	Acc., Pr., Sens., f1-score = .93
[171]	33 PD / 31 HC	Smiling exercise	PD/ non-PD	Relative coordinates and positional jitter for feature extraction (facial expression amplitude and shaking of small facial muscle groups)	LSTM	Pr. = .86, Sens. = .66, f1-score = .75
[100]	107 PD / 1595 HC	Interview process	Hypomimia detection, Med. State	Face detection, Set of hypomimia detection threshold, 72/25	CNN	AUC = .71, Acc. = .63

HO: Hyperparameter Optimization, CV: Cross Validation, LOSO: Leave-One-Subject-Out

diagnosis and support clinical decision-making. Such accurate decision-making systems based on DL models can facilitate early clinical diagnosis of PD and improve patient outcomes.

In this study, we reviewed 87 studies on deep learning for four modalities (gait, upper limb motion, speech, and facial expressions) and their fusion. For the purpose of diagnosing Parkinson's disease and improving the model's accuracy, numerous studies have focused on gait, upper limb movement, and speech signals using both conventional and advanced techniques. However, there may be a need to investigate other diagnostic modalities that can be obtained outside of clinical settings, such as facial muscle movement signals. Additionally, signal fusion from multiple modalities is required to achieve more complex diagnostic goals, such as the disease's severity.

Taking this into consideration, even more difficult is the identification of the disease in its early stages, since the symptoms' heterogeneous appearance and progression result in complicated clinical picture of PD. Therefore, clinical research is also directed at deep phenotyping of the disease, an extensive analysis of the disease's distinct components that exceeds the scope of standard medical records. Deep learning models should be enriched to get high accuracy in the diagnosis of Parkinson's disease. Finally, we believe that metrics other than specificity and sensitivity could be introduced to produce even better provisions for experts in diagnosing PD. These recommendations can probably solve the challenges in enhancing the accuracy of Parkinson's disease classification.

APPENDIX

See Tables VII–XI.

REFERENCES

[1] L. Marsili, G. Rizzo, and C. Colosimo, "Diagnostic criteria for Parkinson's disease: From James Parkinson to the concept of prodromal disease," *Frontiers Neurol.*, vol. 9, p. 156, Mar. 2018.

[2] E. Tolosa, A. Garrido, S. Scholz, and W. Poewe, "Challenges in the diagnosis of Parkinson's disease," *Lancet Neurol.*, vol. 20, no. 5, pp. 385–397, May 2021.

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

[4] H. Tibar et al., "Non-motor symptoms of Parkinson's disease and their impact on quality of life in a cohort of Moroccan patients," *Frontiers Neurol.*, vol. 9, p. 170, Apr. 2018.

[5] K. Kostoglou, K. P. Michmizos, P. Stathis, D. Sakas, K. S. Nikita, and G. D. Mitsis, "Classification and prediction of clinical improvement in deep brain stimulation from intraoperative microelectrode recordings," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 5, pp. 1123–1130, May 2017.

[6] J. Gandhi, A. Gadekar, T. Rajabally, P. Vinayakray-Jani, and D. Ambawade, "Detection of Parkinsons disease via a multi-modal approach," in *Proc. 12th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jul. 2021, pp. 1–7.

[7] Z. Huang et al., "Parkinson's disease classification and clinical score regression via united embedding and sparse learning from longitudinal data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 3357–3371, Aug. 2022.

[8] 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, official J. Movement Disorder Soc.*, vol. 23, no. 15, pp. 2129–2170, 2008, doi: 10.1002/mds.22340.

[9] R. D. Markello, G. Shafiei, C. Tremblay, R. B. Postuma, A. Dagher, and B. Misic, "Multimodal phenotypic axes of Parkinson's disease," *npj Parkinson's Disease*, vol. 7, no. 1, p. 6, 2021.

[10] N. S. Pramod, L. Sajitha, S. Mohanlal, K. Thameem, and S. M. Anzar, "Detection of Parkinson's disease using vocal features: An Eigen approach," in *Proc. 4th Int. Conf. Microelectron., Signals Syst. (ICMSS)*, Nov. 2021, pp. 1–6.

[11] J. Egger et al., "Medical deep learning—A systematic meta-review," *Comput. Methods Programs Biomed.*, vol. 221, Jun. 2022, Art. no. 106874.

[12] A. K. Tyagi and M. M. Nair, *Deep Learning for Clinical and Health Informatics*. Hoboken, NJ, USA: Wiley, 2021, ch. 5, pp. 107–129.

[13] J. Goecks, V. Jalili, L. M. Heiser, and J. W. Gray, "How machine learning will transform biomedicine," *Cell*, vol. 181, no. 1, pp. 92–101, Apr. 2020.

[14] L. Zahid et al., "A spectrogram-based deep feature assisted computer-aided diagnostic system for Parkinson's disease," *IEEE Access*, vol. 8, pp. 35482–35495, 2020.

[15] M. Y. Thanoun and M. T. Yaseen, "A comparative study of Parkinson disease diagnosis in machine learning," in *Proc. 4th Int. Conf. Adv. Artif. Intell.*, Oct. 2020, pp. 23–28.

[16] H. Gunduz, "Deep learning-based Parkinson's disease classification using vocal feature sets," *IEEE Access*, vol. 7, pp. 115540–115551, 2019.

[17] X. Yang, Q. Ye, G. Cai, Y. Wang, and G. Cai, "PD-ResNet for classification of Parkinson's disease from gait," *IEEE J. Transl. Eng. Health Med.*, vol. 10, pp. 1–11, 2022.

[18] J. L. Adams et al., "A real-world study of wearable sensors in Parkinson's disease," *npj Parkinson's Disease*, vol. 7, no. 1, pp. 1–8, Nov. 2021.

[19] S. M. A. Iqbal, I. Mahgoub, E. Du, M. A. Leavitt, and W. Asghar, "Advances in healthcare wearable devices," *npj Flexible Electron.*, vol. 5, no. 1, pp. 1–14, Apr. 2021.

[20] E. Mbunge, B. Muchemwa, S. Jiyane, and J. Batani, "Sensors and healthcare 5.0: Transformative shift in virtual care through emerging digital health technologies," *Global Health J.*, vol. 5, no. 4, pp. 169–177, Dec. 2021.

[21] M. Pistacchi et al., "Gait analysis and clinical correlations in early Parkinson's disease," *Funct. Neurol.*, vol. 32, no. 1, pp. 28–34, 2017.

[22] C. Chatzaki et al., "Can gait features help in differentiating Parkinson's disease medication states and severity levels? A machine learning approach," *Sensors*, vol. 22, no. 24, p. 9937, Dec. 2022.

[23] A. Mirelman et al., "Gait impairments in Parkinson's disease," *Lancet Neurol.*, vol. 18, no. 9, pp. 697–708, Jul. 2019.

[24] Y. A. M. Grimbergen, A. Schrag, G. Mazibrada, G. F. Borm, and B. R. Bloem, "Impact of falls and fear of falling on health-related quality of life in patients with Parkinson's disease," *J. Parkinson's Disease*, vol. 3, no. 3, pp. 409–413, 2013.

[25] K. A. E. Martens et al., "The functional network signature of heterogeneity in freezing of gait," *Brain*, vol. 141, no. 4, pp. 1145–1160, Apr. 2018.

[26] M. Gilat et al., "A systematic review on exercise and training-based interventions for freezing of gait in Parkinson's disease," *npj Parkinson's Disease*, vol. 7, no. 1, pp. 1–18, Sep. 2021.

[27] G. Ebersbach, C. Moreau, F. Gandor, L. Defebvre, and D. Devos, "Clinical syndromes: Parkinsonian gait," *Movement Disorders, Off. J. Movement Disorder Soc.*, vol. 28, no. 11, pp. 1552–1559, Sep. 2013.

[28] P. Tahafchi and J. W. Judy, "Freezing-of-gait detection using wearable-sensor technology and neural-network classifier," *Proc. IEEE SENSORS*, Oct. 2019, pp. 1–4.

[29] J. H. Kordower et al., "Disease duration and the integrity of the nigrostriatal system in Parkinson's disease," *Brain*, vol. 136, pp. 2419–2431, Jul. 2013.

[30] S. A. Schneider, L. Drude, M. Kasten, C. Klein, and J. Hagenah, "A study of subtle motor signs in early Parkinson's disease," *Movement Disorders*, vol. 27, no. 12, pp. 1563–1566, Oct. 2012.

[31] Y. Zhong et al., "A review on pathology, mechanism, and therapy for cerebellum and tremor in Parkinson's disease," *npj Parkinson's Disease*, vol. 8, no. 1, pp. 1–9, Jun. 2022.

[32] E. Rovini, C. Maremmani, and F. Cavallo, "A wearable system to objectify assessment of motor tasks for supporting Parkinson's disease diagnosis," *Sensors*, vol. 20, no. 9, p. 2630, May 2020.

[33] V. Skaramagkas, G. Andrikopoulos, Z. Kefalopoulou, and P. Polychronopoulos, "A study on the essential and Parkinson's arm tremor classification," *Signals*, vol. 2, no. 2, pp. 201–224, Apr. 2021.

[34] J. H. McAuley and C. D. Marsden, "Physiological and pathological tremors and rhythmic central motor control," *Brain*, vol. 123, no. 8, pp. 1545–1567, Aug. 2000.

[35] M. Topka, M. Schneider, C. Zrenner, U. Belardinelli, U. Ziemann, and D. Weiss, "Motor cortex excitability is reduced during freezing of upper limb movement in Parkinson's disease," *npj Parkinson's Disease*, vol. 8, no. 1, pp. 1–7, Nov. 2022.

[36] M. Scholten, A. Schoellmann, A. Ramos-Murguialday, E. López-Larraz, A. Gharabaghi, and D. Weiss, "Transitions between repetitive tapping and upper limb freezing show impaired movement-related beta band modulation," *Clin. Neurophysiol., Off. J. Int. Fed. Clin. Neurophysiol.*, vol. 131, no. 10, pp. 2499–2507, Oct. 2020.

[37] S. T. Nemanich, M. E. McNeely, G. M. Earhart, S. A. Norris, and K. J. Black, "A case of apparent upper-body freezing in Parkinsonism while using a wheelchair," *Frontiers Neurol.*, vol. 8, p. 205, May 2017.

[38] Y. Liu, T. Lee, T. Law, and K. Y. Lee, "Acoustical assessment of voice disorder with continuous speech using ASR posterior features," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 27, no. 6, pp. 1047–1059, Jun. 2019.

[39] B. Suhas et al., "Speech task based automatic classification of ALS and Parkinson's Disease and their severity using log Mel spectrograms," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Jul. 2020, pp. 1–5.

[40] P. Khojasteh, R. Viswanathan, B. Aliahmad, S. Ragnav, P. Zham, and D. K. Kumar, "Parkinson's disease diagnosis based on multivariate deep features of speech signal," in *Proc. IEEE Life Sci. Conf. (LSC)*, Oct. 2018, pp. 187–190.

[41] J. C. V. Correa, T. Arias, J. R. Orozco-Arroyave, and E. Nöth, "A multitask learning approach to assess the dysarthria severity in patients with Parkinson's disease," in *Proc. INTERSPEECH*, Sep. 2018, pp. 456–460.

[42] D. Sztaho, M. G. Tulics, K. Vicsi, and I. Valalik, "Automatic estimation of severity of Parkinson's disease based on speech rhythm related features," in *Proc. 8th IEEE Int. Conf. Cognit. Infocommunications (CogInfoCom)*, Sep. 2017, pp. 11–16.

[43] L. Ricciardi et al., "Hypomimia in Parkinson's disease: An axial sign responsive to levodopa," *Eur. J. Neurol.*, vol. 27, no. 12, pp. 2422–2429, Dec. 2020.

[44] T. Maycas-Cepeda et al., "Hypomimia in Parkinson's disease: What is it telling us?" *Frontiers Neurol.*, vol. 11, p. 6, Jan. 2021.

[45] J. Kimura, "Disorder of interneurons in Parkinsonism. The orbicularis oculi reflex to paired stimuli," *Brain, J. Neurol.*, vol. 96, no. 1, pp. 87–96, 1973.

[46] C. Gasca-Salas and D. Urso, "Association between hypomimia and mild cognitive impairment in de novo Parkinson's disease patients," *Can. J. Neurol. Sci. J. Canadien Des. Sci. Neurol.*, vol. 47, pp. 855–857, Nov. 2020.

[47] R. A. Ramdhani, A. Khojandi, O. Shylo, and B. H. Kopell, "Optimizing clinical assessments in Parkinson's disease through the use of wearable sensors and data driven modeling," *Frontiers Comput. Neurosci.*, vol. 12, p. 72, Sep. 2018.

[48] M. Belić, V. Bobić, M. Badža, N. Šolaja, M. Durić-Jovičić, and V. S. Kostić, "Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—A review," *Clin. Neurol. Neurosurg.*, vol. 184, Sep. 2019, Art. no. 105442.

[49] S. Saravanan et al., "A systematic review of artificial intelligence (AI) based approaches for the diagnosis of Parkinson's disease," *Arch. Comput. Methods Eng.*, vol. 29, pp. 3639–3653, Oct. 2022.

[50] H. W. Loh et al., "Application of deep learning models for automated identification of Parkinson's disease: A review (2011–2021)," *Sensors*, vol. 21, no. 21, p. 7034, Oct. 2021.

[51] J. Mei, C. Desrosiers, and J. Frasnelli, "Machine learning for the diagnosis of Parkinson's disease: A review of literature," *Frontiers Aging Neurosci.*, vol. 13, May 2021, Art. no. 633752.

[52] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ., Keele, U.K., Durham Univ., Durham, U.K., Joint Rep. EBSE 2007-001, 2007.

[53] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," in *Proc. 12th Int. Conf. Eval. Assessment Softw. Eng. (EASE)*, Swindon, U.K., 2008, pp. 68–77.

[54] 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, 2007.

[55] R. Guo, H. Li, C. Zhang, and X. Qian, "A tree-structure-guided graph convolutional network with contrastive learning for the assessment of Parkinsonian hand movements," *Med. Image Anal.*, vol. 81, Oct. 2022, Art. no. 102560.

[56] R. Miao, S. Shokur, A. C. De Lima Pardini, D. Boari, and M. Bouri, "Detecting freezing of gait in Parkinson's disease patient via deep residual network," in *Proc. 20th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2021, pp. 320–325.

[57] L. Tong, J. He, and L. Peng, "CNN-based PD hand tremor detection using inertial sensors," *IEEE Sensors Lett.*, vol. 5, no. 7, pp. 1–4, Jul. 2021.

[58] A. V. A. Shenoy, M. A. Lones, S. L. Smith, and M. Vallejo, "Evaluation of recurrent neural network models for Parkinson's disease classification using drawing data," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 1702–1706.

[59] M. Novotny, T. Tykalova, H. Ruzickova, E. Ruzicka, P. Dusek, and J. Rusz, "Automated video-based assessment of facial bradykinesia in de-novo Parkinson's disease," *npj Digit. Med.*, vol. 5, no. 1, p. 98, Jul. 2022.

[60] L. F. Gomez, A. Morales, J. R. Orozco-Arroyave, R. Daza, and J. Fierrez, "Improving Parkinson detection using dynamic features from evoked expressions in video," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2021, pp. 1562–1570.

[61] P. Varalakshmi, B. T. Priya, A. B. Rithiga, and R. Bhuvaneaswari, "Parkinson disease detection based on speech using various machine learning models and deep learning models," in *Proc. Int. Conf. Syst., Comput., Autom. Netw. (ICSCAN)*, Jul. 2021, pp. 1–6.

[62] J. C. Vasquez-Correa, T. Arias-Vergara, P. Klumpp, P. A. Perez-Toro, J. R. Orozco-Arroyave, and E. Nöth, "End-2-end modeling of speech and gait from patients with Parkinson's disease: Comparison between high quality vs. smartphone data," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 7298–7302.

[63] A. H. Butt, F. Cavallo, C. Maremmani, and E. Rovini, "Biomechanical parameters assessment for the classification of Parkinson disease using bidirectional long short-term memory," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 5761–5764.

[64] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Inf. Softw. Technol.*, vol. 55, no. 12, pp. 2049–2075, Dec. 2013.

[65] T. Jefferson, M. Rudin, S. B. Folse, and F. Davidoff, "Editorial peer review for improving the quality of reports of biomedical studies," *Cochrane Database Systematic Rev.*, vol. 2010, no. 1, Apr. 2007.

[66] N. B. Ali and K. Petersen, "Evaluating strategies for study selection in systematic literature studies," in *Proc. 8th ACM/IEEE Int. Symp. Empirical Softw. Eng. Meas. (ESEM)*, New York, NY, USA: Association for Computing Machinery, Sep. 2014, pp. 1–4.

[67] M. Z. Alom et al., "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, p. 292, Mar. 2019.

[68] Y. H. Bhosale and K. S. Patnaik, "Application of deep learning techniques in diagnosis of COVID-19 (Coronavirus): A systematic review," *Neural Process. Lett.*, pp. 1–53, 2022, doi: 10.1007/s11063-022-11023-0.

[69] R. Sun, Z. Wang, K. E. Martens, and S. Lewis, "Convolutional 3D attention network for video based freezing of gait recognition," in *Proc. Digit. Image Comput., Techn. Appl. (DICTA)*, Dec. 2018, pp. 1–7.

[70] Y. Zhang and D. Gu, "A deep convolutional-recurrent neural network for freezing of gait detection in patients with Parkinson's disease," in *Proc. 12th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2019, pp. 1–6.

[71] N. T. H. Thu and D. S. Han, "Freezing of gait detection using discrete wavelet transform and hybrid deep learning architecture," in *Proc. 12th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Aug. 2021, pp. 448–451.

[72] L. Sigcha et al., "Deep learning approaches for detecting freezing of gait in Parkinson's disease patients through on-body acceleration sensors," *Sensors*, vol. 20, no. 7, p. 1895, Mar. 2020.

[73] S. Mekruksavanich and A. Jitpattanakul, "Detection of freezing of gait in Parkinson's disease by squeeze-and-excitation convolutional neural network with wearable sensors," in *Proc. 15th Int. Conf. Open Source Syst. Technol. (ICOSST)*, Dec. 2021, pp. 1–5.

[74] K. Hu et al., "Graph sequence recurrent neural network for vision-based freezing of gait detection," *IEEE Trans. Image Process.*, vol. 29, pp. 1890–1901, 2020.

[75] N. Naghavi and E. Wade, "Towards real-time prediction of freezing of gait in patients with Parkinson's disease: A novel deep one-class classifier," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, pp. 1726–1736, Apr. 2022.

[76] S. A. Shah et al., "Sensor fusion for identification of freezing of gait episodes using Wi-Fi and radar imaging," *IEEE Sensors J.*, vol. 20, no. 23, pp. 14410–14422, Dec. 2020.

[77] B. B. Ogul and S. Özdemir, "A pairwise deep ranking model for relative assessment of Parkinson's disease patients from gait signals," *IEEE Access*, vol. 10, pp. 6676–6683, 2022.

[78] A. Zhao et al., "Multimodal gait recognition for neurodegenerative diseases," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 9439–9453, Sep. 2022.

[79] G. Paragliola and A. Coronato, "Gait anomaly detection of subjects with Parkinson's disease using a deep time series-based approach," *IEEE Access*, vol. 6, pp. 73280–73292, 2018.

[80] V. Ramesh and E. Bilal, "Detecting motor symptom fluctuations in Parkinson's disease with generative adversarial networks," *npj Digit. Med.*, vol. 5, no. 1, pp. 1–10, Sep. 2022.

[81] V. Skaramagkas, G. Andrikopoulos, Z. Kefalopoulou, and P. Polychronopoulos, "Towards differential diagnosis of essential and Parkinson's tremor via machine learning," in *Proc. 28th Medit. Conf. Control Autom. (MED)*, Sep. 2020, pp. 782–787.

[82] J. J. Hathaliya, H. Modi, R. Gupta, and S. Tanwar, "Deep learning and blockchain-based essential and Parkinson tremor classification scheme," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFO-COM WKSHPS)*, May 2022, pp. 1–6.

[83] S. Shahtalebi, S. F. Atashzar, R. V. Patel, M. S. Jog, and A. Mohammadi, "A deep explainable artificial intelligent framework for neurological disorders discrimination," *Sci. Rep.*, vol. 11, no. 1, Dec. 2021, Art. no. 9630.

[84] C. Taleb, L. Likforman-Sulem, C. Mokbel, and M. Khachab, "Detection of Parkinson's disease from handwriting using deep learning: A comparative study," *Evol. Intell.*, Sep. 2020.

[85] C. Taleb, M. Khachab, C. Mokbel, and L. Likforman-Sulem, "Feature selection for an improved Parkinson's disease identification based on handwriting," in *Proc. 1st IEEE Int. Workshop Arabic Script Anal. Recognit. (ASAR)*, Oct. 2017, pp. 52–56.

[86] Z. Xu, R. Wang, J. Wang, and D. Yu, "Parkinson's disease detection based on spectrogram-deep convolutional generative adversarial network sample augmentation," *IEEE Access*, vol. 8, pp. 206888–206900, 2020.

[87] N. A. Kumar, B. R. Bhatt, P. Anitha, A. K. Yadav, K. K. Devi, and V. C. Joshi, "A new diagnosis using a Parkinson's disease XGBoost and CNN-based classification model using ML techniques," in *Proc. Int. Conf. Adv. Comput. Technol. Appl. (ICACTA)*, Mar. 2022, pp. 1–6.

[88] B. Karan, S. S. Sahu, and K. Mahto, "Stacked auto-encoder based time-frequency features of speech signal for Parkinson disease prediction," in *Proc. Int. Conf. Artif. Intell. Signal Process. (AISP)*, Jan. 2020, pp. 1–4.

[89] L. Moro-Velazquez, J. A. Gomez-Garcia, J. D. Arias-Londoño, N. Dehak, and J. I. Godino-Llorente, "Advances in Parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102418.

[90] L. Ricciardi et al., "Facial emotion recognition and expression in Parkinson's disease: An emotional mirror mechanism?" *PLoS ONE*, vol. 12, no. 1, Jan. 2017, Art. no. e0169110.

[91] J. C. M. Zijlmans, S. E. Daniel, A. J. Hughes, T. Révész, and A. J. Lees, "Clinicopathological investigation of vascular Parkinsonism, including clinical criteria for diagnosis," *Movement Disorders, Off. J. Movement Disorder Soc.*, vol. 19, no. 6, pp. 630–640, Jun. 2004.

[92] L. P. S. Fernández, L. A. S. Pérez, J. J. C. Hernández, and G. de J. Rodríguez Jordán, "Biomechanical signal analysis for evaluation of gait in Parkinson's disease," in *Proc. IEEE 23rd Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, vol. 1, Oct. 2018, pp. 792–799.

[93] C. Fernandes et al., "Discrimination of idiopathic Parkinson's disease and vascular Parkinsonism based on gait time series and the levodopa effect," *J. Biomechanics*, vol. 125, Aug. 2021, Art. no. 110214.

[94] T. Basaklar, Y. Tuncel, and U. Y. Ogras, "Subject-independent freezing of gait (FoG) prediction in Parkinson's disease patients," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2021, pp. 1–6.

[95] G. Shalin, S. Pardoel, J. Nantel, E. D. Lemaire, and J. Kofman, "Prediction of freezing of gait in Parkinson's disease from foot plantar-pressure arrays using a convolutional neural network," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 244–247.

[96] 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 *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2018, pp. 1–8.

[97] B. Karan, S. S. Sahu, and J. R. Orozco-Arroyave, "An investigation about the relationship between dysarthria level of speech and the neurological state of Parkinson's patients," *Biocybernetics Biomed. Eng.*, vol. 42, no. 2, pp. 710–726, Apr. 2022.

[98] B. Valenzuela, J. Arevalo, W. Contreras, and F. Martinez, "A spatio-temporal hypomimic deep descriptor to discriminate parkinsonian patients," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 4192–4195.

[99] A. Abrami, S. Gunzler, C. Kilbane, R. Ostrand, B. Ho, and G. Cecchi, "Automated computer vision assessment of hypomimia in Parkinson disease: Proof-of-principle pilot study," *J. Med. Internet Res.*, vol. 23, no. 2, Feb. 2021, Art. no. e21037.

[100] L. J. Dommershuijsen, A. J. W. Boon, and M. K. Ikram, "Probing the pre-diagnostic phase of Parkinson's disease in population-based studies," *Frontiers Neurol.*, vol. 12, Jul. 2021, Art. no. 702502.

[101] S. Wan, Y. Liang, Y. Zhang, and M. Guizani, "Deep multi-layer perceptron classifier for behavior analysis to estimate Parkinson's disease severity using smartphones," *IEEE Access*, vol. 6, pp. 36825–36833, 2018.

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

[103] J. Prince, F. Andreotti, and M. De Vos, "Multi-source ensemble learning for the remote prediction of Parkinson's disease in the presence of source-wise missing data," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 5, pp. 1402–1411, May 2019.

[104] W. Li, W. Zhu, E. R. Dorsey, and J. Luo, "Predicting Parkinson's disease with multimodal irregularly collected longitudinal smartphone data," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2020, pp. 1106–1111.

[105] S. Wan, Y. Liang, Y. Zhang, and M. Guizani, "Deep multi-layer perceptron classifier for behavior analysis to estimate Parkinson's disease severity using smartphones," *IEEE Access*, vol. 6, pp. 36825–36833, 2018.

[106] M. Ghislieri, V. Agostini, L. Rizzi, M. Knaflitz, and M. Lanotte, "Atypical gait cycles in Parkinson's disease," *Sensors*, vol. 21, no. 15, p. 5079, Jul. 2021.

[107] C. Chatzaki et al., "The smart-insole dataset: Gait analysis using wearable sensors with a focus on elderly and Parkinson's patients," *Sensors*, vol. 21, no. 8, pp. 1–22, 2021.

[108] Z. Kefalopoulou et al., "Pressure sensor insole gait assessment for Parkinson's disease patients: A pilot study," in *Movement Disorders*, vol. 37. Hoboken, NJ, USA: Wiley, 2022, p. S193.

[109] D. Boucharas et al., "Smart insole: A gait analysis monitoring platform targeting Parkinson disease patients based on insoles," Nov. 2022, *arXiv:2212.00109*.

[110] T. Tuncer and S. Dogan, "A novel octopus based Parkinson's disease and gender recognition method using vowels," *Appl. Acoust.*, vol. 155, pp. 75–83, Dec. 2019.

[111] M. A. Hobert, W. Maetzler, K. Aminian, and L. Chiari, "Technical and clinical view on ambulatory assessment in Parkinson's disease," *Acta Neurologica Scandinavica*, vol. 130, no. 3, pp. 139–147, Sep. 2014.

[112] L. C. Benson, A. M. Räisänen, C. A. Clermont, and R. Ferber, "Is this the real life, or is this just laboratory? A scoping review of IMU-based running gait analysis," *Sensors*, vol. 22, no. 5, p. 1722, Feb. 2022.

[113] V. Tsakanikas et al., "Evaluating gait impairment in Parkinson's disease from instrumented insole and IMU sensor data," *Sensors*, vol. 23, no. 8, p. 3902, Apr. 2023.

[114] F. Kluge, H. Gaßner, J. Hannink, C. Pasluosta, J. Klucken, and B. Eskofier, "Towards mobile gait analysis: Concurrent validity and test-retest reliability of an inertial measurement system for the assessment of spatio-temporal gait parameters," *Sensors*, vol. 17, no. 7, p. 1522, Jun. 2017.

[115] F. Bo et al., "IMU-based monitoring for assistive diagnosis and management of IoHT: A review," *Healthcare*, vol. 10, no. 7, p. 1210, Jun. 2022.

[116] Y. Yang, L. Chen, J. Pang, X. Huang, L. Meng, and D. Ming, "Validation of a spatiotemporal gait model using inertial measurement units for early-stage Parkinson's disease detection during turns," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 12, pp. 3591–3600, Dec. 2022.

[117] S. S. Bangera, T. D. Shiyana, G. K. Srinidhi, Y. R. Vasani, M. S. Rao, "MEMS-based IMU for pose estimation," in *Advances in Communication, Signal Processing, VLSI, and Embedded Systems* (Lecture Notes in Electrical Engineering), vol. 614, S. Kalya, M. Kulkarni, and K. Shivaprakasha, Eds. Singapore: Springer, 2020, doi: 10.1007/978-981-15-0626-0_1.

[118] J. Taborri, E. Palermo, S. Rossi, and P. Cappa, "Gait partitioning methods: A systematic review," *Sensors*, vol. 16, no. 1, p. 66, Jan. 2016.

[119] J. Chen et al., "Plantar pressure-based insole gait monitoring techniques for diseases monitoring and analysis: A review," *Adv. Mater. Technol.*, vol. 7, no. 1, Jan. 2022, Art. no. 2100566.

[120] M. Parati et al., "Validation of pressure-sensing insoles in patients with Parkinson's disease during overground walking in single and cognitive dual-task conditions," *Sensors*, vol. 22, no. 17, p. 6392, Aug. 2022.

[121] S. T. Moore, H. G. MacDougall, and W. G. Ondo, "Ambulatory monitoring of freezing of gait in Parkinson's disease," *J. Neurosci. Methods*, vol. 167, no. 2, pp. 340–348, Jan. 2008.

[122] C. Chang, Y. Huang, J. Chen, and C. Lee, "Improving automatic tremor and movement motor disorder severity assessment for Parkinson's disease with deep joint training," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 3408–3411.

[123] P. Y. Chan et al., "Motion characteristics of subclinical tremors in Parkinsonbs disease and normal subjects," *Sci. Rep.*, vol. 12, p. 4021, Mar. 2022.

[124] F. Cordella, A. Paffi, and A. Pallotti, "Classification-based screening of Parkinson's disease patients through voice signal," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2021, pp. 1–6.

[125] Y. Campos-Roca, F. Calle-Alonso, C. J. Perez, and L. Naranjo, "Computational diagnosis of Parkinson's disease from speech based on regularization methods," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 1127–1131.

[126] X. Liu et al., "A freezing of gait detection system based on stepwise feature extraction method," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Aug. 2020, pp. 2060–2065.

[127] L. Martínez-Villaseñor, H. Ponce, and J. Nuñez-Martinez, "Artificial hydrocarbon networks for freezing of gait detection in Parkinso's disease," in *Proc. Int. Conf. Mechatronics, Electron. Automat. Eng. (ICMEAE)*, Nov. 2020, pp. 3–8.

[128] B. Shi, S. C. Yen, A. Tay, D. M. L. Tan, N. S. Y. Chia, and W. L. Au, "Convolutional neural network for freezing of gait detection leveraging the continuous wavelet transform on lower extremities wearable sensors data," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 5410–5415.

[129] V. G. Torvi, A. Bhattacharya, and S. Chakraborty, "Deep domain adaptation to predict freezing of gait in patients with Parkinson's disease," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 1001–1006.

[130] L. Borzì, G. Olmo, C. A. Artusi, and L. Lopiano, "Detection of freezing of gait in people with Parkinson's disease using smartphones," in *Proc. IEEE 44th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*, Jul. 2020, pp. 625–635.

[131] A.-M. Tautan, A.-G. Andrei, and B. Ionescu, "Freezing of gait detection for Parkinson's disease patients using accelerometer data: Case study," in *Proc. Int. Conf. e-Health Bioeng. (EHB)*, Dec. 2020, pp. 1–4.

[132] L. Borzì, L. Sigcha, D. Rodríguez-Martín, and G. Olmo, "Real-time detection of freezing of gait in Parkinson's disease using multi-head convolutional neural networks and a single inertial sensor," *Artif. Intell. Med.*, vol. 135, Jan. 2023, Art. no. 102459.

[133] H. Gokul, P. Suresh, B. H. Vignesh, R. P. Kumaar, and V. Vijayaraghavan, "Gait recovery system for Parkinson's disease using machine learning on embedded platforms," in *Proc. 14th IEEE Int. Syst. Conf. (SysCon)*, Aug. 2020, pp. 1–8.

[134] A. Prado, K. Kwei, N. Vanegas-Arroyave, and S. K. Agrawal, "Identification of freezing of gait in Parkinson's patients using instrumented shoes and artificial neural networks," in *Proc. 8th IEEE RAS/EMBS Int. Conf. Biomed. Robot. Biomechatronics (BioRob)*, Nov. 2020, pp. 68–73.

[135] G. Shalin, S. Pardoel, E. D. Lemaire, J. Nantel, and J. Kofman, "Prediction and detection of freezing of gait in Parkinson's disease from plantar pressure data using long short-term memory neural-networks," *J. Neuroeng. Rehabil.*, vol. 18, no. 1, pp. 1–15, Dec. 2021.

[136] N. Naghavi, A. Miller, and E. Wade, "Towards real-time prediction of freezing of gait in patients with Parkinson's disease: Addressing the class imbalance problem," *Sensors*, vol. 19, no. 18, p. 3898, Sep. 2019.

[137] A. Salimi-Badr and M. Hashemi, "A neural-based approach to aid early Parkinson's disease diagnosis," in *Proc. 11th Int. Conf. Inf. Knowl. Technol. (IKT)*, Dec. 2020, pp. 23–25.

[138] P. Ren et al., "Gait rhythm fluctuation analysis for neurodegenerative diseases by empirical mode decomposition," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 52–60, Jan. 2017.

[139] Y. Liu, B. Oubre, C. Duval, S. I. Lee, and J. Daneault, "A kinematic data-driven approach to differentiate involuntary choreic movements in individuals with neurological conditions," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 12, pp. 3784–3791, Dec. 2022.

[140] W. Zeng, C. Yuan, Q. Wang, F. Liu, and Y. Wang, "Classification of gait patterns between patients with Parkinson's disease and healthy controls using phase space reconstruction (PSR), empirical mode decomposition (EMD) and neural networks," *Neural Netw., Off. J. Int. Neural Netw. Soc.*, vol. 111, pp. 64–76, Mar. 2019.

[141] N. S. Hoang, Y. Cai, C. Lee, Y. O. Yang, C. Chui, and M. C. H. Chua, "Gait classification for Parkinson's disease using stacked 2D and 1D convolutional neural network," in *Proc. Int. Conf. Adv. Technol. Commun. (ATC)*, Oct. 2019, pp. 44–49.

[142] A. S. Alharthi, A. J. Casson, and K. B. Ozanyan, "Gait spatiotemporal signal analysis for Parkinson's disease detection and severity rating," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1838–1848, Jan. 2021.

[143] J. F. Reyes, J. S. Montealegre, Y. J. Castano, C. Urcuqui, and A. Navarro, "LSTM and convolution networks exploration for Parkinson's diagnosis," in *Proc. IEEE Colombian Conf. Commun. Comput. (COLCOM)*, Jun. 2019, pp. 1–4.

[144] L. C. Guayacán, E. Rangel, and F. Martínez, "Towards understanding spatio-temporal parkinsonian patterns from salient regions of a 3D convolutional network," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 3688–3691.

[145] L. Aversano, M. L. Bernardi, M. Cimitile, and R. Pecori, "Early detection of Parkinson disease using deep neural networks on gait dynamics," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–8.

[146] T. Varrecchia et al., "An artificial neural network approach to detect presence and severity of Parkinson's disease via gait parameters," *PLoS ONE*, vol. 16, no. 2, Feb. 2021, Art. no. e0244396.

[147] I. El Maachi, G.-A. Bilodeau, and W. Bouachir, "Deep 1D-convnet for accurate Parkinson disease detection and severity prediction from gait," *Expert Syst. Appl.*, vol. 143, Apr. 2020, Art. no. 113075.

[148] X. Cao et al., "Video based shuffling step detection for parkinsonian patients using 3D convolution," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 641–649, 2021.

[149] C. Lin, F. Wang, T. Kuo, P. Huang, S. Chen, and L. Fu, "Early detection of Parkinson's disease by neural network models," *IEEE Access*, vol. 10, pp. 19033–19044, 2022.

[150] S. Moon et al., "Classification of Parkinson's disease and essential tremor based on balance and gait characteristics from wearable motion sensors via machine learning techniques: A data-driven approach," *J. Neuroeng. Rehabil.*, vol. 17, no. 1, Dec. 2020, Art. no. 125.

[151] R. Kaur, R. W. Motl, R. Sowers, and M. E. Hernandez, "A vision-based framework for predicting multiple sclerosis and Parkinson's disease gait dysfunctions—A deep learning approach," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 1, pp. 190–201, Jan. 2023.

[152] L. Fraiwan and O. Hassanin, "Computer-aided identification of degenerative neuromuscular diseases based on gait dynamics and ensemble decision tree classifiers," *PLoS ONE*, vol. 16, no. 6, Jun. 2021, Art. no. e0252380.

[153] H. Thummikarat and P. Chongstitvatana, "An implementation of machine learning for Parkinson's disease diagnosis," in *Proc. 18th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol. (ECTI-CON)*, May 2021, pp. 258–261.

[154] R. Castrillon et al., "Characterization of the handwriting skills as a biomarker for Parkinson's disease," in *Proc. 14th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2019, pp. 1–5.

[155] O. Bazgir, J. Frounchi, S. A. H. Habibi, L. Palma, and P. Pierleoni, "A neural network system for diagnosis and assessment of tremor in Parkinson disease patients," in *Proc. 22nd Iranian Conf. Biomed. Eng. (ICBME)*, Nov. 2015, pp. 1–5.

[156] X. Xing, N. Luo, S. Li, L. Zhou, C. Song, and J. Liu, "Identification and classification of parkinsonian and essential tremors for diagnosis using machine learning algorithms," *Frontiers Neurosci.*, vol. 16, Mar. 2022.

[157] H. Wu, Y. Zhang, X. Wu, and F. Yang, "Assessment of upper limb tremors in patients with Parkinson's disease based on displacement and acceleration information," in *Proc. 5th Int. Conf. Autom., Control Robot. Eng. (CACRE)*, Sep. 2020, pp. 177–182.

[158] M. Sun et al., "TremorSense: Tremor detection for Parkinson's disease using convolutional neural network," in *Proc. IEEE/ACM Conf. Connected Health, Appl., Syst. Eng. Technol. (CHASE)*, Dec. 2021, pp. 1–10.

[159] H. Zhang, A. Wang, D. Li, and W. Xu, "DeepVoice: A voiceprint-based mobile health framework for Parkinson's disease identification," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Mar. 2018, pp. 214–217.

[160] J. P. Folador et al., "On the use of histograms of oriented gradients for tremor detection from sinusoidal and spiral handwritten drawings of people with Parkinson's disease," *Med. Biol. Eng. Comput.*, vol. 59, no. 1, pp. 195–214, Jan. 2021.

[161] H. B. Kim et al., "Wrist sensor-based tremor severity quantification in Parkinson's disease using convolutional neural network," *Comput. Biol. Med.*, vol. 95, pp. 140–146, Apr. 2018.

[162] A. Papadopoulos, K. Kyritsis, S. Bostanjopoulou, L. Klingelhoefer, R. K. Chaudhuri, and A. Delopoulos, "Multiple-instance learning for in-the-wild Parkinsonian tremor detection," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 6188–6191.

[163] M. D. Hssayeni, J. Jimenez-Shahed, M. A. Burack, and B. Ghoraani, "Wearable sensors for estimation of Parkinsonian tremor severity during free body movements," *Sensors*, vol. 19, no. 19, p. 4215, Sep. 2019.

[164] E. J. Alqahtani, F. H. Alshamrani, H. F. Syed, and S. O. Olatunji, "Classification of Parkinson's disease using NNge classification algorithm," in *Proc. 21st Saudi Comput. Soc. Nat. Comput. Conf. (NCC)*, Apr. 2018, pp. 1–7.

[165] M. Anila and G. Pradeepini, "Diagnosis of Parkinson's disease using deep neural network model," in *Proc. Int. Conf. Smart Gener. Comput., Commun. Netw. (SMART GENCON)*, Oct. 2021, pp. 1–7.

[166] F. Demir, A. Sengur, A. Ari, K. Siddique, and M. Alswaitti, "Feature mapping and deep long short term memory network-based efficient approach for Parkinson's disease diagnosis," *IEEE Access*, vol. 9, pp. 149456–149464, 2021.

[167] M. B. Shafeena and S. Vijayan, "Parkinson's disease prognosis using the ResNet-50 model from speech features," in *Proc. Int. Conf. Innov. Sci. Technol. Sustain. Develop. (ICISTSD)*, Aug. 2022, pp. 282–286.

[168] P. Deepa and R. Khilar, "Parkinson's disease classification from speech signal parameters using deep neural network," in *Proc. IEEE 4th Int. Conf. Cybern., Cognition Mach. Learn. Appl. (ICCCMLA)*, Oct. 2022, pp. 44–48.

[169] M. Munsif, M. Ullah, B. Ahmad, M. Sajjad, and F. A. Cheikh, "Monitoring neurological disorder patients via deep learning based facial expressions analysis," in *Proc. IFIP Int. Conf. Artif. Intell. Appl. Innov.*, in IFIP Advances in Information and Communication Technology, vol. 652, 2022, pp. 412–423.

[170] B. Jin, Y. Qu, L. Zhang, and Z. Gao, "Diagnosing Parkinson disease through facial expression recognition: Video analysis," *J. Med. Internet Res.*, vol. 22, no. 7, Jul. 2020, Art. no. e18697.