algorithms

MDPI

Review

Parkinson's Disease Detection Using Machine Learning Algorithms: A Comprehensive Review

Jelica Cincović [1], Miloš Cvetanović [1], Milica Djurić-Jovičić [2], Nebojsa Bacanin [3,4,*] and Boško Nikolić [1]

[1] School of Electrical Engineering, University of Belgrade, 11000 Belgrade, Serbia; jelica@etf.bg.ac.rs (J.C.); cmilos@etf.bg.ac.rs (M.C.); nbosko@etf.bg.ac.rs (B.N.)
[2] Innovation Center of the School of Electrical Engineering in Belgrade, 11000 Belgrade, Serbia; milica.djuric@etf.bg.ac.rs
[3] Faculty of Informatics and Computing, Singidunum University, 11000 Belgrade, Serbia
[4] Saveetha School of Engineering, SIMATS, Thandalam, Chennai 602105, Tamilnadu, India
* Correspondence: nbacanin@singidunum.ac.rs

Abstract

Parkinson's disease (PD) is a progressive neurodegenerative disorder in which early detection remains a major clinical challenge due to heterogeneous motor and non-motor manifestations and the lack of reliable biomarkers. In recent years, machine learning (ML) and deep learning (DL) methods have been increasingly investigated as decision-support tools for PD screening using diverse clinical and behavioral data. This review synthesizes PD detection studies published between 2017 and 2025, systematically analyzing 32 representative works across multiple modalities, including MRI, PET, EEG, REM sleep biomarkers, voice recordings, gait signals, handwriting/drawing tasks, and finger-tapping measurements. Across the reviewed literature, high classification performance is frequently reported, with CNN-based and hybrid DL architectures achieving particularly strong results in imaging and time-series settings, while classical ML approaches such as SVM and ensemble models remain competitive for engineered feature-based datasets. However, the review also reveals major barriers to reliable translation, including small datasets, inconsistent evaluation protocols, limited external validation, and the risk of performance inflation caused by non-subject-independent data splitting. Overall, this review provides a structured and modality-oriented reference of algorithms, datasets, and performance trends, while highlighting key methodological gaps and practical priorities for developing robust and clinically deployable PD detection systems.

Keywords: Parkinson's disease; machine learning; deep learning; feature selection; hyperparameter optimization

Check for updates

Academic Editor: Iman Dehzangi

Received: 28 December 2025
Revised: 27 February 2026
Accepted: 2 March 2026
Published: 4 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The brain, predominantly composed of neurons and glial cells, facilitates information processing through synaptic connections, with neurotransmitters such as dopamine, serotonin, and histamine mediating signal transmission [1]. Dopamine, synthesized in regions including the substantia nigra, plays a central role in motor control, reward processing, sleep regulation and many other processes in the human body [2]. The degeneration of dopaminergic neurons within the substantia nigra results in dopamine depletion, which constitutes a core pathological feature of Parkinson's disease (PD).

PD manifests through a spectrum of motor symptoms, including tremors, bradykinesia, rigidity, and postural and gait disturbances, as well as non-motor manifestations such as cognitive impairment, sleep dysfunction, and psychiatric symptoms [3]. The heterogeneity of clinical presentation, combined with contributions from additional neurotransmitter systems, genetic predispositions, and environmental factors, poses significant challenges for early and accurate diagnosis. Conventional diagnostic approaches, including magnetic resonance imaging (MRI), positron emission tomography (PET), and electroencephalography (EEG), as well as clinical evaluation of motor function, although informative, are often time-intensive, subjective, and susceptible to inter-observer variability. Current therapeutic strategies are symptomatic, with Levodopa being the most widely administered pharmacological treatment to restore dopamine levels and alleviate motor deficits.

In this context, machine learning (ML) represents a promising strategy to enhance PD detection by identifying subtle and complex patterns within multimodal datasets that may be overlooked by human assessment. Supervised ML algorithms, trained on labeled datasets, can perform accurate classification of patients using data derived from neuroimaging, electrophysiological recordings, speech analysis, gait assessment, and fine motor evaluations. Deep learning approaches, notably convolutional neural networks (CNNs) and hybrid architectures, further optimize performance through automated feature extraction and modeling of spatial-temporal dependencies. Such data-driven methodologies have the potential to complement conventional diagnostics, reduce clinical subjectivity, and facilitate earlier intervention, thereby improving patient management and outcomes.

This paper explores the algorithms employed in supervised learning and deep learning used for solving classification problems, which were found applicable in detecting PD in the recent literature. The algorithms that are explored are the Naïve Bayes classifier (NB), K-Nearest Neighbors (KNN), and Support Vector Machines (SVM). Additionally, a group of decision tree-based algorithms are covered, including Decision Trees (DT), Random Forest (RF), and boosting algorithms such as AdaBoost (AB), Gradient Boosting (GB), Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) [4–10]. Several deep learning algorithms are also discussed, including Artificial Neural Networks (ANN), Feedforward Neural Networks (FNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) [11–14].

The main contributions of this paper are as follows:

- A comprehensive literature review focused on the use of machine learning techniques for the early detection of PD,
- The identification and comparison of commonly used supervised learning algorithms in PD classification tasks,
- Suggestions for future research directions aimed at improving the accuracy, robustness, and clinical applicability of machine learning models in this domain.

The paper is organized into chapters, which are outlined in the sections that follow. The Section 2 reviews the datasets that were used. The Section 3 provides an overview of the literature, divided by datasets and used algorithms. The Section 4 provides a detailed analysis of the results, possible improvements and future directions. Finally, a conclusion is given in the Section 5.

2. Datasets

The data that goes into machine learning algorithms has a big impact on the algorithm's results. Different data types can be collected to predict PD, and depending on the observed symptoms can be divided into motor and non-motor symptoms. Motor symptoms include gait and hand movements (handwriting, drawing, and finger-tapping), while non-motor symptoms include voice and medical data (MRI, EEG, etc.).

In this chapter, an overview of the datasets (DS) used in the literature is provided (Table 1), and for each dataset the following is stated: the source of the data, the type of data collected, the method of data collection, the total number of people in the dataset, the number of healthy people (HC) in the dataset, the number of people with PD in the dataset and a link to the dataset if it is publicly available. These key details are crucial for understanding the scope, relevance, and potential applications of the dataset in clinical studies and machine learning model development.

The datasets presented herein were sourced from the literature, with a substantial number being publicly accessible through the hyperlinks provided in the table. Datasets lacking associated links are available upon request, and inquiries should be directed to the respective dataset providers. For each dataset, the total number of participating patients is indicated, along with the counts of PD patients and healthy controls (HC). In datasets involving medical data analysis, each patient is usually represented by a single recording, whereas non-medical datasets include multiple recordings per patient.

The analysis includes a total of 26 datasets, of which 8 are medical datasets comprising 3 MRI, 3 EEG, 1 REM and 1 PET dataset. The remaining 18 datasets cover non-medical modalities, including 6 datasets based on voice recordings, 3 on gait analyses, 5 on spiral wave drawing, 2 on hand movement and finger-tapping recordings and 2 hybrid datasets combining multiple modalities. Among these, 14 datasets analyze non-motor symptoms, and 12 motor symptoms. Dataset sizes vary significantly, ranging from as few as 5 to as many as 2042 participants. Notably, several datasets exceed 500 participants, providing substantial resources for large-scale studies. Among these, 14 datasets are publicly available via direct links, facilitating open access and reproducibility. The medical datasets emphasize neuroimaging and electrophysiological data that are critical for clinical research, while the non-medical datasets offer diverse behavioral and motor function analyses. This variety and size distribution present a comprehensive resource for research across multiple domains in this field.

Table 1. Datasets review.

Label	Source	Type of Data	Collection Method	Total Number	HC Number	PD Number	Link
DS.1.	University Medical Center Groningen, University of Genoa, Clinic University of Navarra	PET scan (non-motor)	All images were spatially normalized onto an 18F-FDG-PET template in Montreal Neurological Institute (MNI) brain space29 using SPM12 v6 software	310	145	165	
DS.2.	Parkinson's Progression Markers Initiative	3T T1-weighted MRI scans (non-motor)	All scans were taken with the same type of scanner, and were collected in 20–30 min field of view (FoV) of all the scans including vertex, cerebellum, and pons.	406	203	203	
DS.3.	Parkinson's Progression Markers Initiative	T2-weighted MRI scans (non-motor)	/	500	250	250	
DS.4.	Department of Neurology, National Institute of Mental Health and Neurosciences (NIMHANS)	Neuromelanin-sensitive MRI scans (non-motor)	3D neuromelanin contrast-sensitive sequences were acquired using 50 slices, with an acquisition time of 4 min 12.9 s. T1, T2, fluid attenuation inversion recovery and diffusion-weighted images were collected in order to eliminate the possibility of brain lesions	100	35	45 + 20 atypical parkinsonian disorders	
DS.5.	PRED-CT data repository	Electroencephalogram (EEG) signals (non-motor)	The data was collected for On and Off medication patients, with a rate of 500 samples per second using sintered Ag/AgCl electrodes for 64 channels. Patients first came to obtain control under therapy, and the second visit occurred at least 15 h after therapy.	50	25	25	https://predict.cs.unm.edu/ accessed on 27 February 2026

Table 1. *Cont.*

Label	Source	Type of Data	Collection Method	Total Number	HC Number	PD Number	Link
DS.6.	University of San Diego, California	Electroencephalogram (EEG) signals (non-motor)	3 min recordings with 512 Hz frequency and 32-channel Biosemi ActiveTwo EEG system.	31	16	15	https://openneuro.org/datasets/ds002778/versions/1.0.5 accessed on 27 February 2026
DS.7.	Henan Provincial People's Hospital;	Electroencephalogram (EEG) signals (non-motor)	5 min recordings with 128 Hz frequency and 14 channel emotive EPOC neuroheadset.	40	20	20	
DS.8.	Parkinson's Progression Markers Initiative	Rapid Eye Movement (REM) and olfactory loss, cerebrospinal fluid data (CFD), and dopaminergic imaging markers (non-motor)	/	584	183	401	
DS.9.	PC-GITA	Voice recordings (non-motor)	People speaking Spanish were recorded pronouncing vowels, words and monologs with normal intonation.	100	50	50	
DS.10.	Saarbruecken Voice Database	Voice recordings (non-motor)	German voice recordings of vowels with normal, low, high and rising–falling intonation with 50 Hz frequency.	2042	687	1355	
DS.11.	Open Data Commons	Voice recordings (non-motor)	Vowel recording with no information about PD.	1676	/	/	https://www.kaggle.com/datasets/darubiano57/dataset-of-vowels accessed on 27 February 2026
DS.12.	UCI public repository	Voice recordings (non-motor)	Patients had telemonitoring devices at home for 6 months, and 200 recordings with 16 voice features were recorded for each patient.	42	/	42	https://archive.ics.uci.edu/dataset/189/parkinsons+telemonitoring accessed on 27 February 2026

Table 1. *Cont.*

Label	Source	Type of Data	Collection Method	Total Number	HC Number	PD Number	Link
DS.13.	University of Oxford, in collaboration with the National Centre for Voice and Speech	Voice recordings (non-motor)	Voice measures such as jitter, shimmer, etc., were recorded for each individual.	31	8	23	https://www.kaggle.com/datasets/pypiahmad/oxford-parkinsons-disease-detection-dataset accessed on 27 February 2026
DS.14.	Department of Neurology in Cerrahpasa Faculty of Medicine, Istanbul University	Voice recordings (non-motor)	Individuals were recorded pronouncing the vowel "a" three times, in front of a microphone with a set frequency of 44.1 Hz.	252	64	188	https://archive.ics.uci.edu/dataset/470/parkinson+s+disease+classification accessed on 27 February 2026
DS.15.	Physionet	Gait recordings (motor)	Each individual wore 8 sensors under the foot, and the vertical ground reaction force was measured at a frequency of 100 Hz while they walked at their usual pace.	166	73	93	https://physionet.org/content/gaitpdb/1.0.0/ accessed on 27 February 2026
DS.16.	Department of Medico-Surgical Sciences and Biotechnologies, University of Rome Sapienza	Gait recordings (motor)	The initial dataset consists of 22 features collected through the inertial measurement unit (IMU) sensor.	161	80	81	
DS.17.	REMPARK project	Gait recordings (motor)	Waist-placed IMU sensors sampled data at 200 Hz while individuals completed some basic house activities such as walking from sofa to bed, standing up and sitting down, or showing off the apartment.	21	/	21	

Table 1. *Cont.*

Label	Source	Type of Data	Collection Method	Total Number	HC Number	PD Number	Link
DS.18.	Department of Neurology in Cerrahpasa Faculty of Medicine, Istanbul University	Spiral drawings (motor)	Individuals used a Wacom Cintiq 12WX tablet and drew certain drawings on the screen with a digital pen. Information about coordinates, pressure and grip angle was stored for each drawing.	77	15	62	https://archive.ics.uci.edu/dataset/395/parkinson+disease+spiral+drawings+using+digitized+graphics+tablet accessed on 27 February 2026
DS.19.	Faculty of Electrical Engineering and Communication, Brno University of Technology	Spiral drawings (motor)	Drawings were taken using Intuos 4M pen of frequency 200 Hz.	75	37	38	https://bdalab.utko.fekt.vut.cz/ accessed on 27 February 2026
DS.20.	Faculty of Sciences-UNESP-Bauru	Spiral and meander drawings (motor)	Individuals filled out a form that included both spiral and meander drawings.	92	18	74	https://wwwp.fc.unesp.br/~papa/pub/datasets/Handpd/ accessed on 27 February 2026
DS.21.	Kaggle	Spiral and meander drawings (motor).	Individuals were assigned four different writing tasks, including circle, spiral, meander, and signals.	66	35	31	https://wwwp.fc.unesp.br/~papa/pub/datasets/Handpd/ accessed on 27 February 2026
DS.22.	Kaggle	Spiral and wave drawings (motor)	/	102	51	51	https://www.kaggle.com/datasets/kmader/parkinsons-drawings/data accessed on 27 February 2026

Table 1. *Cont.*

Label	Source	Type of Data	Collection Method	Total Number	HC Number	PD Number	Link
DS.23.	Federal State Budget Scientific Institution "Scientific Center of Neurology" and Scientific and Educational Medical and Technological Center (SEMTCN) of BMSTU	Hand movement (motor)	Individuals completed a few motion tasks, such as finger-tapping, pronation–supination of the hands, and opening–closing hand movements with the maximum possible amplitude and speed. The recordings lasted 16 s for each task and the LeapMotion sensor was used.	32	16	16	
DS.24.	Movement Disorders Unit at the Clinic of Neurology, Belgrade	Finger-tapping (motor)	Sensors attached to the thumb and index finger of the right hand of the individual recorded six times individuals alternately tapped the thumb and index finger for 15 s, with a 1 min break.	54	11	14 + 19 atypical parkinsonian disorders	https://github.com/innovation-center-etf/FINGER-TAPPING-PARKINSONISMS-DATABASE accessed on 27 February 2026
DS.25.	Johns Hopkins academic movement disorder clinic	Voice, posture, gait, finger-tapping data (hybrid)	Individuals recorded several tests using the LG Optimus S mobile phone four times a day, which included a voice test, posture test, gait test, finger-tapping test and reaction time test.	20	10	10	
DS.26.	University of Pittsburgh, Carnegie Mellon University and University of Washington	Voice, gait, finger-tapping data (hybrid)	Individuals wore Pebble watch that was connected to a smartphone.	5	/	/	

3. Literature Review

This chapter presents an overview of the literature on PD detection using the previously described machine learning algorithms and datasets.

This systematic review was conducted following PRISMA guidelines to ensure transparency and reproducibility in selecting studies on PD detection using machine learning algorithms.

Identification: A structured search was performed across Google Scholar, PubMed, and IEEE Xplore using keywords related to PD (e.g., "Parkinson's disease", "PD", "Parkinsonism"), machine learning (e.g., "machine learning", "deep learning", "artificial intelligence"), and early detection (e.g., "early detection", "early diagnosis", "classification"). The initial search returned over 1000 articles.

Screening: Titles and abstracts were reviewed, and studies not relevant to PD detection or not applying machine learning were excluded. Duplicates were also removed, resulting in approximately 140 articles for full-text assessment.

Eligibility: Full-text articles were evaluated based on the following inclusion criteria:

- Use of machine learning or deep learning algorithms for PD detection;
- Published between 2017 and 2025;
- At least one algorithm in the study reported a classification accuracy $\geq$80%.

The $\geq$80% accuracy threshold was applied to focus on studies demonstrating strong predictive performance for at least one algorithm, while acknowledging that some studies reported lower accuracy for other models within the same work. Methodologically important studies with lower-performing algorithms were also considered qualitatively if they provided significant insights or novel approaches relevant to PD detection. Additionally, a number of studies were found to use the same datasets, with very similar algorithms and comparable results. To avoid redundancy and the overrepresentation of nearly identical approaches, the most representative or best-performing study was selected in such cases. This strategy allowed the review to emphasize a broader diversity of methods and modeling approaches applied to PD detection.

Inclusion: After applying these criteria, 32 studies were included in the final review. These studies form the basis of the literature analysis presented in this chapter.

Each paper (in Tables 2–5) states which dataset was used and from which year it originates, which group the analyzed symptoms belong to, which machine learning algorithms and metrics were used, and what results were achieved.

Based on the analysis of these tables, the goal is to identify the most effective models for non-motor and motor symptom detection across different data modalities by systematically comparing their performance. This is important because it highlights consistent trends across studies and helps guide future researchers toward the most promising methods. A detailed analysis of these tables is provided in Section 5.

In the reviewed literature, a total of 32 studies were identified, including 4 based on MRI data, 2 on EEG, 1 on REM sleep data, 1 on PET scans, 10 on voice recordings, 9 on gait analysis, and 5 on hand movement data and drawings. Among the datasets, DS.15. and DS.9. were the most frequently used across multiple studies, indicating their popularity and relevance within the field. A range of machine learning models was used, with SVM being the most common. Other frequent methods included KNN, RF, CNNs, hybrid CNN + LSTM and ResNet + LSTM, as well as ensemble techniques like XGBoost, LightGBM, and AB. Traditional models such as DTs, NB, and ANNs were also applied. Some studies focused on model optimization and tuning, while others explored the performance of pre-trained CNNs. The diversity of approaches highlights ongoing efforts to improve diagnostic accuracy across different data types and modalities.

Figure 1 illustrates the distribution of publications between journals and conferences, while Figure 2 presents the most prominent publishers represented in the reviewed literature.

Table 2. Literature review of papers analyzing medical data.

Paper	Dataset	Year	Symptom Types	Algorithms	Metrics	Results
[15]	DS.1.	2019.	Non-motor (PET)	CNN	Accuracy AUC	86% 0.94
[16]	DS.2.	2018.	Non-motor (MRI)	KNN	Accuracy Sensitivity Specificity	85% 77.5% 92.5%
				NB	Accuracy Sensitivity Specificity	87.5% 82.5% 92.5%
				SVM	Accuracy Sensitivity Specificity	90% 87.5% 92.5%
[17]	DS.2.	2020.	Non-motor (MRI)	CNN	Accuracy Recall Precision AUC	95.29% 0.943 0.927 0.98
[18]	DS.3.	2018.	Non-motor (MRI)	CNN	Accuracy AUC	96% 0.995
[19]	DS.4.	2019.	Non-motor (MRI)	CNN	Accuracy AUC	85.7% 0.911
[20]	DS.5.	2020.	Non-motor (EEG)	CNN + LSTM	Accuracy F1-score Cohan Kappa	99.22% 0.994 0.984
[21]	DS.6. DS.7.	2021.	Non-motor (EEG)	CNN	Accuracy Precision F1-score	99.7% $\pm$ 0.093 0.99 $\pm$ 0.01 0.99 $\pm$ 0.0009
[22]	DS.8.	2020.	Non-motor (REM, CFD)	FNN	Accuracy AUC F1-score	96.60% 0.98 0.97
				RF	Accuracy AUC F1-score	95.66% 0.98 0.96
				KNN	Accuracy AUC F1-score	95.16% 0.97 0.96
				SVM	Accuracy AUC F1-score	95.96% 0.98 0.97

Table 3. Literature review of papers analyzing voice data.

Paper	Dataset	Year	Symptom Types	Algorithms	Metrics	Results
[23]	DS.9.	2019.	Non-motor (voice)	CNN	Precision F1-score	0.92 0.92
[24]	DS.9.	2021.	Non-motor (voice)	ResNet + LSTM	Accuracy Precision	97.20% 0.97
[25]	DS.9.	2022.	Non-motor (voice)	SVM	Accuracy Sensitivity Specificity	84.3% 84.0% 84.7%
[26]	DS.9.	2022.	Non-motor (voice)	CNN	Accuracy AUC	99% 0.89
	DS.10. DS.11.		used for tuning			
[27]	DS.12.	2021.	Non-motor (voice)	RNN + LSTM	Accuracy Precision F1-score	95.8% 0.92 0.96
	DS.13.				Accuracy Precision F1-score	82.2% 0.82 0.90
[28]	DS.13.	2025.	Non-motor (voice)	SVM	Accuracy Precision Recall F1-score	92.3% 0.91 1 0.95
				RF	Accuracy Precision Recall F1-score	89.76% 0.91 0.96 0.93
				DT	Accuracy Precision Recall F1-score	87.17% 0.9 0.93 0.92
[29]	DS.13.	2025.	Non-motor (voice)	NB	Accuracy AUC	69.3% 0.86
				DT	Accuracy AUC	75.2% 0.79
				RF	Accuracy AUC	85.2% 0.93
				KNN	Accuracy AUC	60.6% 0.80
[30]	DS.14.	2020.	Non-motor (voice)	XGBoost	Accuracy	85.60%

Table 3. *Cont.*

Paper	Dataset	Year	Symptom Types	Algorithms	Metrics	Results
[31]	DS.14.	2021.	Non-motor (voice)	GB	Accuracy Precision F1-score	92.95% 0.95 0.95
				LightGB	Accuracy Precision F1-score	93.39% 0.95 0.95
				XGboost	Accuracy Precision F1-score	92.07% 0.93 0.95
				AB	Accuracy Precision F1-score	87.22% 0.91 0.92
[32]	DS.14.	2022.	Non-motor (voice)	KNN	Accuracy Precision F1-score	95.4% 0.952 0.955

Table 4. Literature review of papers analyzing gait data.

Paper	Dataset	Year	Symptom Types	Algorithms	Metrics	Results
[33]	DS.15.	2017.	Motor (gait)	RF	Accuracy Cohan Kappa F1-score	98.04% 0.95 0.98
[34]	DS.15.	2018.	Motor (gait)	SVM	Accuracy AUC	94.8% 0.97
[35]	DS.15.	2020.	Motor (gait)	NB	Accuracy F1-score	69.7% 0.70
				DT	Accuracy F1-score	99.4% 0.99
				SVM	Accuracy F1-score	97.6% 0.97
[36]	DS.15.	2020.	Motor (gait)	CNN	Accuracy	98.7%
[37]	DS.15.	2021.	Motor (gait)	CNN	Accuracy	96.52%
[38]	DS.15.	2025.	Motor (gait)	SVM	Accuracy F1-score	98% 0.977
				KNN	Accuracy F1-score	97% 0.968
				DT	Accuracy F1-score	97% 0.967
				ANN	Accuracy F1-score	95% 0.948
				LSTM	Accuracy F1-score	96% 0.956
				CNN	Accuracy F1-score	99.1% 0.999

Table 4. *Cont.*

Paper	Dataset	Year	Symptom Types	Algorithms	Metrics	Results
[39]	DS.15.	2025.	Motor (gait)	CNN	Accuracy	98.6%
[40]	DS.16.	2022.	Motor (gait)	SVM	Accuracy F1-score AUC	86% 0.85 0.85
				RF	Accuracy F1-score AUC	82% 0.82 0.82
				DT	Accuracy F1-score AUC	79% 0.80 0.79
				KNN	Accuracy F1-score AUC	74% 0.74 0.74
				ANN	Accuracy F1-score AUC	77% 0.78 0.77
[41]	DS.17.	2018.	Motor (gait)	CNN	Accuracy	89%

Table 5. Literature review of papers analyzing hand movement data.

Paper	Dataset	Year	Symptom Types	Algorithms	Metrics	Results
[42]	DS.18.	2018.	Motor (hand movement)	CNN	Accuracy	88%
[43]	DS.18.	2019.	Motor (hand movement)	CNN	Accuracy F1-score AUC	96.5% 0.97 0.99
[44]	DS.19.	2021.	Motor (hand movement)	Pre-trained CNNs	Accuracy	62%
	DS.20.				Accuracy	90%
	DS.21.				Accuracy	90.41%
	DS.22.				Accuracy	98.31%
[45]	DS.23.	2020.	Motor (hand movement)	KNN	Accuracy	87.5%
				SVM	Accuracy	89.1%
				DT	Accuracy	82.8%
				RF	Accuracy	94.5%
[46]	DS.24.	2023.	Motor (hand movement)	KNN	Accuracy	85.15%

Papers in Table 2 demonstrate that machine learning and deep learning models achieve high performance in non-motor disease classification using PET, MRI, EEG, and sleep-related datasets. CNNs outperform traditional classifiers, with accuracies ranging from 85% to over 99% and AUC values reaching up to 0.995. Hybrid models such as CNN-LSTM show exceptional results on EEG data, achieving accuracies above 99%.

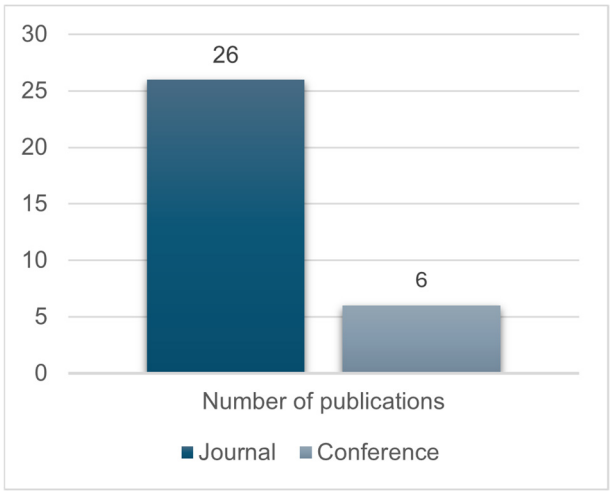

Figure 1. Distribution of publications between journals and conferences.

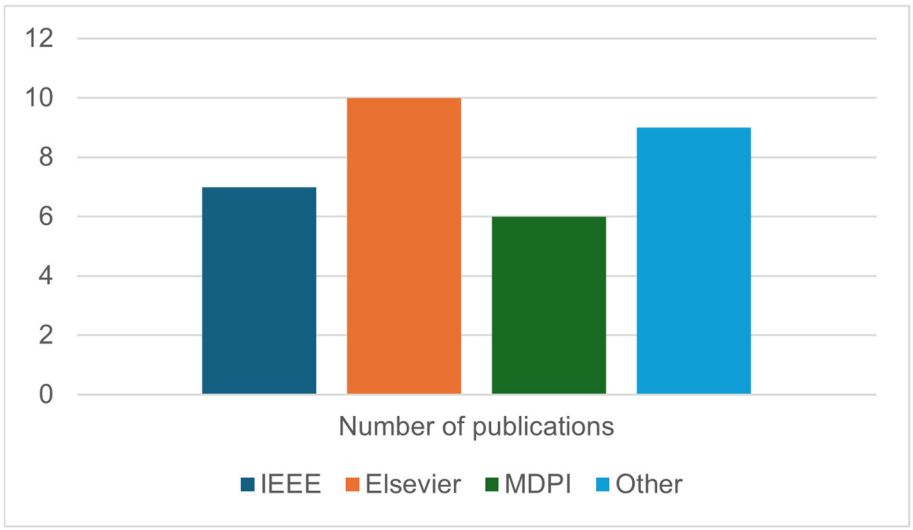

Figure 2. Distribution of publishers over included publications.

Papers in Table 3, on non-motor voice-based datasets show that deep learning models generally outperform basic machine learning approaches. CNN-based and hybrid architectures (such as ResNet + LSTM and RNN + LSTM) achieve high accuracies reaching up to 99% and strong precision and F1-scores, indicating robust feature extraction from voice signals. Classical models like SVM, RF and GB also achieve high performance, with accuracies exceeding 90% in several cases.

Papers in Table 4, on motor symptom detection using gait data indicate strong performance across both traditional machine learning and deep learning approaches. Classical models such as RF, SVM, and DT achieve high accuracy above 97%. Deep learning models, particularly CNNs, consistently provide superior and more stable results, reaching accuracies above 99% in recent studies.

Papers in Table 5, on motor symptom analysis using hand movement data and finger-tapping tests show varying performance depending on the model and dataset. CNN-based approaches generally achieve strong results, with accuracies ranging from 88% to over 96% and AUC values reaching 0.99. Pre-trained CNNs exhibit mixed performance across multiple datasets, indicating sensitivity to data quality and domain adaptation. Traditional machine learning models, particularly RF and SVM, also demonstrate competitive accuracy, exceeding 90% in some cases.

4. Discussion

In this chapter, the discussion focuses on the analysis of the relevant literature and the identification of potential improvements. Key findings from previous studies are evaluated, with particular attention to the instructions for choosing the most appropriate model for PD detection. Based on this analysis, suggestions for future enhancements and research directions are proposed.

4.1. Strengths and Weaknesses of Algorithms

Machine learning algorithms are an important tool in the early detection of PD, because of their ability to recognize subtle patterns in complex datasets. Although these algorithms offer significant advantages, they face certain limitations, including a dependence on data quality and potential bias. Understanding both the strengths and weaknesses (shown in Table 6) of these approaches is crucial to their effective application in clinical practice.

Table 6. Strengths and weaknesses of machine learning algorithms in PD detection.

Algorithms	Strengths	Weaknesses
NB	Simple to implement Fast to train	Assumes independence between features Sensitive to imbalanced datasets
KNN	Simple and intuitive No assumptions about data distribution	Slow for large datasets Sensitive to outliers Requires careful selection of parameters
SVM	Effective for high-dimensional data Handles nonlinear relationships between features	Slow for large datasets Requires careful parameter tuning
DT	Easy to interpret and visualize Requires minimal data preprocessing	Prone to overfitting Sensitive to noisy data
RF	Reduces overfitting compared to single decision trees	Slow for large datasets Hard to interpret
AB	Combines multiple weak learners Handles complex relationships between features	Sensitive to noisy data and outliers Requires careful parameter tuning
GB	Achieves high predictive accuracy Handles heterogeneous data	Computationally intensive Requires careful parameter tuning
XGBoost	Highly accurate and efficient Supports regularization to reduce overfitting Handle missing values and complex nonlinear feature interactions	Computationally intensive for very large datasets Requires careful parameter tuning
LightGB	Extremely fast and memory-efficient Can handle large datasets and high-dimensional features efficiently	Sensitive to noisy data Requires careful parameter tuning
FNN	Can model complex nonlinear relationships Flexible architecture	Requires a relatively large amount of data Prone to overfitting
RNN	Can capture temporal dependencies in shorter sequential data Flexible for various time-series tasks	Vanishing or exploding gradient Computationally expensive
LSTM	Captures long-term dependencies in sequential data Can model temporal patterns in PD progression	Computationally expensive Requires careful parameter tuning
CNN	Particularly effective for image data Captures spatial patterns and local structures in data	Requires large datasets to avoid overfitting Computationally expensive

4.2. Literature Analysis

In the process of machine learning, the first step is the collection and analysis of data. First, it is very important to decide which data should be collected, both in terms of usefulness for the result and in terms of simplicity for the patients themselves. The analyzed symptoms that represent the dataset's data can be divided into motor and non-motor. Analysis of non-motor symptoms, such as head scans, has produced reliable results; however, it can be costly and uncomfortable for the patient. Voice data collection is simpler, and the results are not lacking. Analysis of motor symptoms, such as gait or hand movement, also proved to be a solid basis and certainly a simpler option for the patient. Therefore, the aim is to create datasets with as much data as possible that are easily accessible to the user. The papers in Table 2 are published between 2018 and 2021, and they analyze non-motor symptoms using PET-, MRI-, EEG-, and REM-related datasets. CNNs are the dominant approach and demonstrate consistently high classification performance across all data modalities, with reported accuracies above 85% and AUC values exceeding 0.90. MRI-based CNN models achieve accuracies of up to 96% and AUC values reaching 0.995. EEG-based methods report the highest performance metrics, specifically CNN-LSTM architectures with accuracies above 99%, accompanied by high F1-scores and Cohen's Kappa values. Conventional machine learning algorithms, including KNN, SVM, RF and NB, achieve competitive results on REM-related datasets, with accuracies exceeding 95%, but remain inferior to deep learning approaches. The reported results demonstrate that CNN-based models provide superior performance for non-motor symptom classification, with measurable variation across data modalities.

Papers in Table 3 are published between 2019 and 2025, and report the use of voice-based datasets for non-motor symptom classification. The evaluated deep learning models include CNN, RNN-LSTM, and ResNet-LSTM architectures, with reported accuracy values ranging from 95.8% to 99% and F1-scores up to 0.96. Classical machine learning algorithms, including SVM, RF, DT, NB and KNN, report accuracy values ranging from 60.6% to 95.4%, with corresponding variations in precision, recall, F1-score, and AUC. Ensemble methods such as GB, LightGB, XGBoost, and AB report accuracies between 87.22% and 93.39%. The reported results indicate that deep learning and ensemble-based approaches achieve higher and more consistent performance than standalone traditional classifiers for voice-based non-motor symptom classification.

The papers in Table 4 are published between 2017 and 2025 and report the use of gait-based datasets for motor symptom classification. The evaluated algorithms include RF, SVM, NB, DT, KNN, ANN, CNN, and LSTM. Reported accuracy values range from 69.7% to 99.4%, with F1-scores reaching up to 0.999 and AUC values up to 0.97. Classical machine learning methods show heterogeneous results. NB reports the lowest accuracy of 69.7%, while DT and SVM achieve accuracies above 97% on the same dataset. Deep learning models, particularly CNN-based approaches, report consistently high performance, with accuracies between 89% and 99.1% across multiple studies and years. LSTM and ANN models report accuracy values between 95% and 96%. These results reflect a consistent advantage of CNN-based models for gait-based motor symptom classification, with performance remaining dependent on dataset characteristics.

The papers in Table 5 are published between 2018 and 2023 and report the use of hand movement datasets for motor symptom classification. The evaluated algorithms include CNN, pre-trained CNNs, KNN, SVM, DT and RF. Reported accuracy values range from 62% to 98.31%, with F1-score up to 0.97 and AUC up to 0.99. CNN-based models report high performance, with accuracies between 88% and 96.5% across multiple datasets and years. Pre-trained CNNs show variable results, with one study reporting only 62% accuracy, while other implementations achieve 90% or higher. Classical machine learning methods

demonstrate heterogeneous performance, with RF achieving 94.5% accuracy, SVM 89.1%, KNN between 85.15% and 87.5%, and DT 82.8%. These results suggest that CNN-based approaches provide the most consistent performance for hand movement-based motor symptom classification, while classical and pre-trained models exhibit greater variability across datasets.

Across all reviewed papers, both motor and non-motor symptom detection show high potential using machine learning and deep learning techniques applied to diverse data types such as MRI, EEG, voice, gait, and hand movement signals. Deep learning models, especially CNNs and hybrid architectures consistently achieve superior performance, exceeding 95% accuracy, while traditional machine learning methods remain competitive particularly when robust feature engineering and preprocessing are performed on the data. Overall, the evidence confirms that data-driven models are effective tools for the accurate and early detection of neurodegenerative diseases such as PD.

Table 7 presents a comparative summary of the top three algorithm categories, including their best representative models, average performance, key advantages and limitations, and applicability across different data modalities for motor and non-motor symptom detection in PD.

Table 7. Comparison of top-performing algorithms across data modalities.

Rank	Algorithm Category	Best Representative	Performance	Advantages	Limitations	Data Types and Scenarios
1	**CNN-based Deep Learning**	**CNN/CNN + LSTM**	Accuracy: 95–99.7% F1: 0.95–0.994 AUC: 0.89–0.995	High accuracy Spatial–temporal modeling Automated feature extraction	Requires large datasets High computational cost Lower interpretability	MRI, PET, EEG, voice, gait, hand movement Motor and non-motor symptom detection
2	**Ensemble Methods**	**Random Forest** (*others*: *GB*, *LightGB*, *XGBoost*)	Accuracy: 89–98% F1: 0.93–0.98 AUC: 0.93–0.98	Noise-robust Dataset-stable Interpretable	Needs feature engineering (for boosting methods) Slower on large datasets	Voice, REM, gait, hand movement Small-to-medium datasets Baseline or exploratory analyses
3	**Classical ML**	**SVM** (*others*: *KNN*, *NB*, *DT*)	Accuracy: 60–95% F1: 0.7–0.97 AUC: 0.79–0.97	Simple and fast Effective on small/medium datasets	Sensitive to noise Requires feature engineering	MRI, REM, hand movement, gait Small-to-medium datasets When interpretability is critical

Additionally, Figure 3 shows instructions for choosing the most appropriate model (accuracies over 90%) based on collected data, with references to the corresponding literature.

In addition to highlighting the best-performing algorithms and studies, it is equally important to consider the risk of overfitting when interpreting reported results. A reliable assessment of machine learning models in PD detection requires careful attention to evaluation protocols. Key components include train–test splits, cross-validation schemes, and reporting of confidence intervals or other statistical measures, which collectively ensure that reported performance reflects true generalization rather than overfitting.

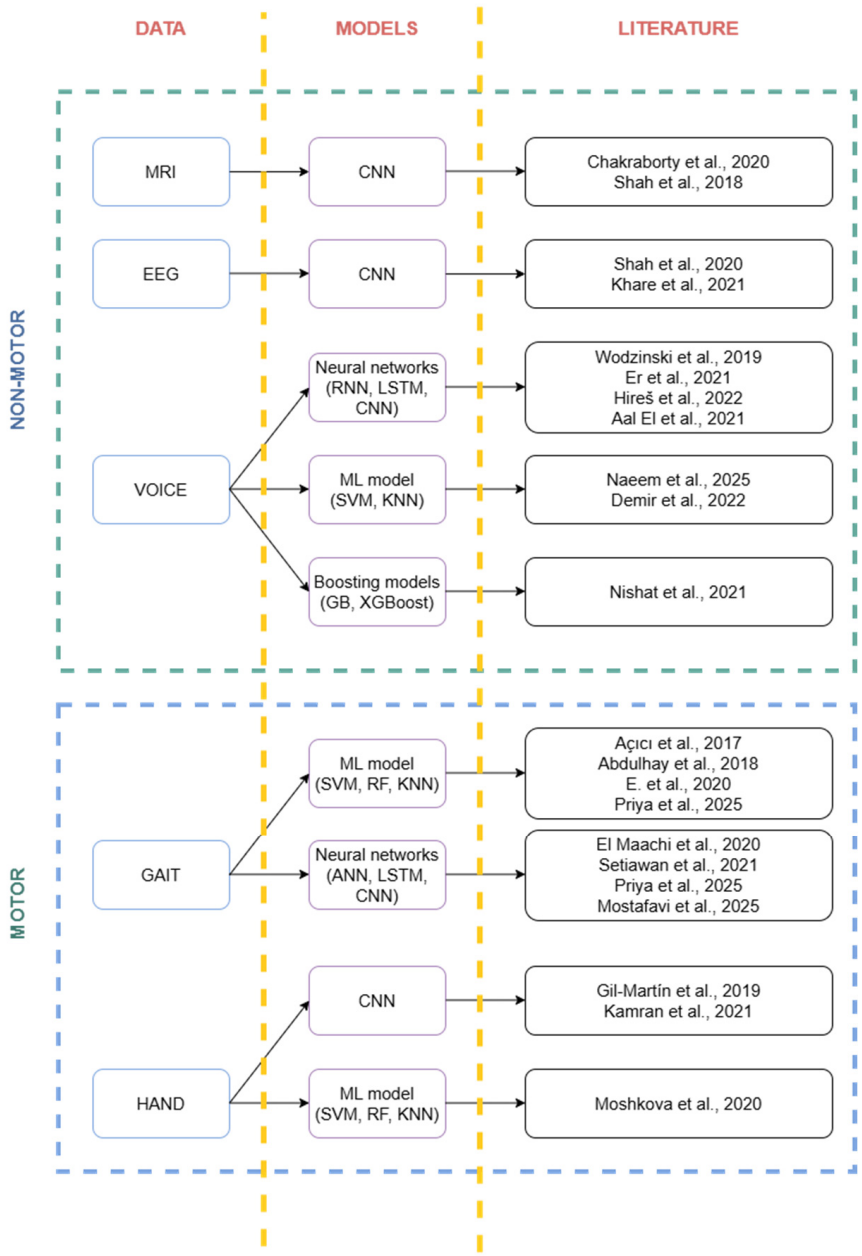

Figure 3. Instructions for choosing the most appropriate model for PD detection ([17,18,20,21,23,24,26–28,31–39,43–45]).

PD datasets are often relatively small and heterogeneous, and may contain multiple recordings per participant (e.g., voice, gait, handwriting, EEG). Without appropriate validation, models risk learning subject-specific patterns rather than disease-related features, which can inflate reported accuracies. While many studies did not explicitly report these evaluation details, their inclusion is essential for accurate and reproducible performance assessment. Methods such as leave-one-subject-out cross-validation or nested cross-validation are particularly valuable in this context, as they separate data at the patient level and provide more realistic estimates of predictive performance.

In addition, reporting confidence intervals or variability measures around performance metrics allows for a better understanding of the uncertainty and stability of the models, facilitating meaningful comparisons across studies and modalities. Transparent and standardized evaluation practices are therefore crucial for developing robust, reproducible, and clinically applicable PD detection models.

4.3. Methodological Challenges and Future Research Directions

After collecting the data and selecting features that contribute the most to the solution, it is important to optimize the classification models. This involves fine-tuning the model's parameters to achieve the best possible performance. By carefully adjusting hyperparameters such as regularization coefficients, learning rate or tree depth, the model can be improved to identify complex patterns more effectively. As many of these hyperparameters exist, choosing the best combination becomes a very difficult problem.

Recent studies reviewed in Tables 2–5 indicate that model performance is highly dependent on appropriate hyperparameter configuration, particularly for deep learning and ensemble methods. Consequently, systematic optimization strategies have become an important research focus in PD detection. Metaheuristics are powerful optimization techniques that aim to find the optimal solutions as simply and efficiently as possible. Drawing inspiration from diverse fields such as physics, mathematics, and nature, these algorithms mimic processes found in the natural world. Although metaheuristics do not guarantee global optimality, they enable the efficient exploration of large and complex search spaces, which is particularly relevant for high-dimensional PD datasets.

In the field of PD detection, metaheuristics can be used in the process of feature selection, as in References [47–50], or in the process of optimizing model hyperparameters, which has recently emerged as a strategy to improve robustness and generalization across datasets. In Reference [51] (2022), hyperparameter optimization is performed with the help of a rat swarm optimizer (RSO). Three datasets were used, and different KNN, RF, and DF algorithms were analyzed. Reference [52] (2024) uses a finger-tapping dataset (DS.24.) and compares the RNN and XGBoost algorithms with different metaheuristics for hyperparameter optimization. Metaheuristics include genetic algorithms, particle swarm optimization, firefly algorithm, whale optimization algorithm, brainstorm optimization, COLSHADE algorithm and sinh cosh optimizer. Obtained accuracies exceed 90%. Reference [53] (2025) uses gait dataset (DS.15.) and optimizes LSTM with the PSO metaheuristic. Accuracy is around 89%.

In addition, metaheuristics not only improve the performance of the models but also offer flexibility across different datasets and types of algorithms. However, as observed in several reviewed studies, high reported accuracies are often obtained on single datasets without external validation. Despite reported accuracies frequently exceeding 95% across different modalities (Tables 2–5), most studies rely on single-dataset evaluation, limiting conclusions about their clinical robustness and real-world deployment. Therefore, future research should emphasize cross-dataset validation, independent test cohorts, and standardized evaluation protocols to assess their true generalizability.

Furthermore, the substantial variability in dataset size (ranging from fewer than 30 to more than 2000 participants) complicates the direct comparison of reported results and highlights the urgent need for standardized benchmarking frameworks and reproducible evaluation pipelines. By optimizing hyperparameters with metaheuristics, it is possible to explore search spaces more effectively than with traditional methods. Nevertheless, the increased computational complexity and potential risk of overfitting require careful experimental design, especially when dealing with relatively small PD cohorts.

Another direction that could be investigated in more detail is the potential combination and examination of different types of symptoms at the same time during diagnosis, as was achieved in References [54,55]. Both papers analyze voice, gait and finger-tapping data simultaneously through smartphone apps. Reference [54] uses DS.25. and Reference [55] uses DS.26. Although these studies demonstrated the feasibility of multimodal remote monitoring, they were conducted on small cohorts and limited hardware configurations. Based on the analysis presented in Tables 3–5, individual modalities (voice, gait, hand

movement) already achieve high classification performance when analyzed separately. However, systematic fusion strategies combining motor and non-motor biomarkers remain insufficiently explored in large-scale, longitudinal settings. Future research should therefore focus on multimodal feature fusion techniques, domain adaptation methods, and the longitudinal modeling of disease progression.

From a translational perspective, an important future direction is the integration of validated machine learning models into clinical decision-support systems. However, before deployment in healthcare institutions, several challenges must be addressed, including model interpretability, regulatory compliance, reproducibility across devices, and integration with existing clinical workflows.

Although smartphone-based and wearable technologies provide a highly accessible platform for data acquisition, their clinical reliability depends on the standardization of data collection procedures and the ability of models to generalize across heterogeneous recording environments.

Smartphone-based PD monitoring has the potential to combine multiple motor and non-motor modalities, including voice recordings, inertial gait measurements, finger-tapping dynamics, and drawing trajectories. As demonstrated by the reviewed literature (Tables 3–5), these individual modalities already achieve strong classification performance when analyzed independently. Therefore, a key future research challenge is not only to implement multimodal acquisition, but to systematically evaluate whether multimodal fusion provides statistically significant improvements over single-modality models and whether such improvements remain stable across different populations, devices, and clinical stages. Future studies should investigate optimal fusion strategies, including the early fusion of raw signals, feature-level fusion, and decision-level ensemble approaches.

Another critical limitation of smartphone-based monitoring is the variability introduced by real-world acquisition. Differences in sensor specifications, sampling frequencies, microphone quality, background noise, lighting conditions, and user behavior may substantially affect the extracted features and degrade model robustness. Furthermore, patient adherence and inconsistent recording conditions can result in incomplete datasets and missing values, requiring reliable preprocessing pipelines and noise-resilient learning strategies. Consequently, future research should prioritize cross-device reproducibility, domain adaptation methods, and the robust handling of missing or corrupted data in order to improve generalizability.

Longitudinal monitoring represents another major opportunity for smartphone and wearable systems, as PD is a progressive disorder characterized by gradual symptom evolution. Continuous acquisition enables temporal modeling approaches that may support the early detection of subtle deterioration patterns and individualized disease tracking. However, longitudinal analysis introduces additional confounders, including medication effects, fatigue, circadian variability, and short-term symptom fluctuations. Therefore, future work should explore time-series models and personalized learning frameworks capable of distinguishing disease progression from transient changes. This is particularly relevant for early-stage PD, where differences between healthy controls and patients may be subtle and difficult to distinguish from single-session measurements.

Finally, privacy and ethical aspects are essential for the scalable deployment of smartphone-based systems, as they require the collection of sensitive medical and behavioral data. Privacy-preserving and distributed learning strategies, such as federated learning, may represent a promising direction for enabling multi-center collaborations without centralized data sharing. Overall, while smartphone-based PD detection offers strong potential for accessible screening and monitoring, its clinical translation requires rigorous validation on large cohorts, standardized benchmarking protocols, and prospective

real-world evaluation studies to bridge the gap between experimental performance and deployable medical solutions.

5. Conclusions

Reviewing the existing literature is essential for understanding the progress and challenges in PD detection, as well as for identifying opportunities for improvement. Numerous studies [56–60] have been published on this topic over the past decade, demonstrating a strong research interest in early and accurate PD diagnosis. This paper reviews the literature in the field of detection of PD in the period from 2017 to 2025. The symptoms of PD are various, but they can be roughly divided into motor and non-motor symptoms. For non-motor symptoms, studies predominantly analyze MRI and EEG data, as well as voice recordings. When combined with CNN-based and hybrid deep learning architectures, several studies report classification accuracies exceeding 99% for specific datasets. In the case of motor symptoms, gait and hand movements (drawings, finger-tapping) are analyzed the most, and in addition to CNNs, which also perform well here, SVM, KNN and decision trees and algorithms based on these methods stand out. Reported accuracies in these modalities frequently exceed 95%, with some studies reporting values above 98%. The use of metaheuristics contributes to the improvement in the results. Based on this, CNNs demonstrate excellent performance for both types of PD symptoms, although careful consideration of dataset size, cohort diversity, validation protocols, and appropriate regularization remains essential to avoid overfitting and inflated performance estimates.

Future research directions include the multimodal integration of motor and non-motor biomarkers, standardized evaluation frameworks, and robust cross-dataset validation. Mobile and wearable technologies may provide a scalable infrastructure for continuous and non-invasive symptom monitoring; however, their clinical deployment requires large-scale validation, reproducible benchmarking, and the careful consideration of interpretability, privacy, and regulatory requirements.

Equally important are efforts to raise public awareness about the value of prevention and early detection—not only for PD but also for a wide range of neurodegenerative and chronic illnesses. Through educational campaigns and integration with healthcare systems, this approach can contribute to earlier diagnoses, reduce the burden on healthcare systems, and improve quality of life for many individuals.

Overall, while machine learning and deep learning approaches demonstrate strong potential for accurate PD detection, their broader clinical adoption depends on methodological rigor, multi-center validation, and translational alignment between computational advances and real-world healthcare practice.

Author Contributions: The manuscript was developed collaboratively by all authors, who participated in scientific discussions, critical revisions, and final approval of the submitted version. The specific contributions are outlined below: Conceptualization: M.C. and M.D.-J. Methodology and literature structuring: M.C., B.N. and N.B. Systematic literature review and data curation: J.C. and M.C. Formal analysis and comparative evaluation of algorithms and datasets: J.C. and B.N. Biomedical domain interpretation and clinical contextualization: M.D.-J. Writing: J.C., M.D.-J., N.B., M.C. and B.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia under contract number 451-03-33/2026-03/200223.

Data Availability Statement: The datasets analyzed in this study are those referenced in the literature and used in the cited works. Links to publicly available repositories for these datasets are provided in Section 3, while for datasets that are not publicly accessible, access may be requested directly from the corresponding authors of the respective publications.

Acknowledgments: The research was partially conducted in the premises of the Palace of Science, Miodrag Kostić Endowment–Centre for Applied Artificial Intelligence.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PD	Parkinson's disease
MRI	Magnetic resonance imaging
PET	Positron emission tomography
EEG	Electroencephalographic scans
REM	Rapid eye movement
NB	Naïve bayes classifier
KNN	K-nearest neighbors
SVM	Support Vector Machines
DT	Decision trees
RF	Random forest
AB	Ada boosting
GB	Gradient boosting
XGBoost	Extreme gradient boosting
LightGB	Light gradient boosting
ANN	Artificial neural networks
FNN	Feed-forward neural networks
RNN	Recurrent neural networks
LSTM	Long short-term memory
CNN	Convolutional neural networks
DS	Dataset
HC	Healthy control
RSO	Rat swarm optimizer
PSO	Particle swarm optimizer

References

1. Hyman, S.E. Neurotransmitters. *Curr. Biol.* **2005**, *15*, R154–R158. [CrossRef]
2. Iversen, S.D.; Iversen, L.L. Dopamine: 50 Years in Perspective. *Trends Neurosci.* **2007**, *30*, 188–193. [CrossRef]
3. Kalia, L.V.; Lang, A.E. Parkinson's Disease. *Lancet* **2015**, *386*, 896–912. [CrossRef]
4. Berrar, D. Bayes' Theorem and Naive Bayes Classifier. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*; Elsevier: Amsterdam, The Netherlands, 2018; Volume 1–3, pp. 403–412.
5. Guo, G.; Wang, H.; Bell, D.; Bi, Y.; Greer, K. KNN Model-Based Approach in Classification. In *On the Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2003; Volume 2888, pp. 986–996.
6. Pisner, D.A.; Schnyer, D.M. Support Vector Machine. In *Machine Learning: Methods and Applications to Brain Disorders*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 101–121.
7. Kotsiantis, S.B. Decision Trees: A Recent Overview. *Artif. Intell. Rev.* **2013**, *39*, 261–283. [CrossRef]
8. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
9. Mayr, A.; Binder, H.; Gefeller, O.; Schmid, M. The Evolution of Boosting Algorithms: From Machine Learning to Statistical Modelling. *Methods Inf. Med.* **2014**, *53*, 419–427. [CrossRef] [PubMed]
10. Fan, J.; Ma, X.; Wu, L.; Zhang, F.; Yu, X.; Zeng, W. Light Gradient Boosting Machine: An Efficient Soft Computing Model for Estimating Daily Reference Evapotranspiration with Local and External Meteorological Data. *Agric. Water Manag.* **2019**, *225*, 105758. [CrossRef]
11. Dongare, A.D.; Kharde, R.R.; Kachare, A.D. Introduction to Artificial Neural Network. *Int. J. Eng. Innov. Technol.* **2012**, *2*, 189–194.
12. Marhon, S.A.; Cameron, C.J.F.; Kremer, S.C. Recurrent Neural Networks. In *Handbook on Neural Information Processing*; Intelligent Systems Reference Library; Springer: Berlin/Heidelberg, Germany, 2013; Volume 49, pp. 29–65. [CrossRef]

13. Graves, A. Long Short-Term Memory. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 385, pp. 37–45.
14. O'Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. *arXiv* **2015**, arXiv:1511.08458. [CrossRef]
15. Manzanera, O.M.; Meles, S.K.; Leenders, K.L.; Renken, R.J.; Pagani, M.; Arnaldi, D.; Nobili, F.; Obeso, J.; Oroz, M.R.; Morbelli, S.; et al. Scaled Subprofile Modeling and Convolutional Neural Networks for the Identification of Parkinson's Disease in 3D Nuclear Imaging Data. *Int. J. Neural Syst.* **2019**, *29*, 1950010. [CrossRef]
16. Cigdem, O.; Demirel, H. Performance Analysis of Different Classification Algorithms Using Different Feature Selection Methods on Parkinson's Disease Detection. *J. Neurosci. Methods* **2018**, *309*, 81–90. [CrossRef] [PubMed]
17. Chakraborty, S.; Aich, S.; Kim, H.C. Detection of Parkinson's Disease from 3T T1 Weighted MRI Scans Using 3D Convolutional Neural Network. *Diagnostics* **2020**, *10*, 402. [CrossRef] [PubMed]
18. Shah, P.M.; Zeb, A.; Shafi, U.; Zaidi, S.F.A.; Shah, M.A. Detection of Parkinson Disease in Brain MRI Using Convolutional Neural Network. In *2018 24th International Conference on Automation and Computing (ICAC)*; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
19. Shinde, S.; Prasad, S.; Saboo, Y.; Kaushick, R.; Saini, J.; Pal, P.K.; Ingalhalikar, M. Predictive Markers for Parkinson's Disease Using Deep Neural Nets on Neuromelanin Sensitive MRI. *NeuroImage Clin.* **2019**, *22*, 101748. [CrossRef] [PubMed]
20. Shah, S.A.A.; Zhang, L.; Bais, A. Dynamical System Based Compact Deep Hybrid Network for Classification of Parkinson Disease Related EEG Signals. *Neural Netw.* **2020**, *130*, 75–84. [CrossRef]
21. Khare, S.K.; Bajaj, V.; Acharya, U.R. PDCNNet: An Automatic Framework for the Detection of Parkinson's Disease Using EEG Signals. *IEEE Sens. J.* **2021**, *21*, 17017–17024. [CrossRef]
22. Wang, W.; Lee, J.; Harrou, F.; Sun, Y. Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning. *IEEE Access* **2020**, *8*, 147635–147646. [CrossRef]
23. Wodzinski, M.; Skalski, A.; Hemmerling, D.; Orozco-Arroyave, J.R.; Noth, E. Deep Learning Approach to Parkinson's Disease Detection Using Voice Recordings and Convolutional Neural Network Dedicated to Image Classification. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*; IEEE: Piscataway, NJ, USA, 2019; pp. 717–720.
24. Er, M.B.; Isik, E.; Isik, I. Parkinson's Detection Based on Combined CNN and LSTM Using Enhanced Speech Signals with Variational Mode Decomposition. *Biomed. Signal Process. Control.* **2021**, *70*, 103006. [CrossRef]
25. Pah, N.D.; Motin, M.A.; Kumar, D.K. Phonemes Based Detection of Parkinson's Disease for Telehealth Applications. *Sci. Rep.* **2022**, *12*, 9687. [CrossRef]
26. Hireš, M.; Gazda, M.; Drotár, P.; Pah, N.D.; Motin, M.A.; Kumar, D.K. Convolutional Neural Network Ensemble for Parkinson's Disease Detection from Voice Recordings. *Comput. Biol. Med.* **2022**, *141*, 105021. [CrossRef]
27. El Aal, H.A.A.; Taie, S.A.; El-Bendary, N. An Optimized Rnn-Lstm Approach for Parkinson's Disease Early Detection Using Speech Features. *Bull. Electr. Eng. Inform.* **2021**, *10*, 2503–2512. [CrossRef]
28. Naeem, I.; Ditta, A.; Mazhar, T.; Anwar, M.; Saeed, M.M.; Hamam, H. Voice Biomarkers as Prognostic Indicators for Parkinson's Disease Using Machine Learning Techniques. *Sci. Rep.* **2025**, *15*, 12129. [CrossRef]
29. Xu, H.; Xie, W.; Pang, M.; Li, Y.; Jin, L.; Huang, F.; Shao, X. Non-Invasive Detection of Parkinson's Disease Based on Speech Analysis and Interpretable Machine Learning. *Front. Aging Neurosci.* **2025**, *17*, 1586273. [CrossRef]
30. Abdurrahman, G.; Sintawati, M. Implementation of Xgboost for Classification of Parkinson's Disease. *J. Phys. Conf. Ser.* **2020**, *1538*, 012024. [CrossRef]
31. Nishat, M.M.; Hasan, T.; Nasrullah, S.M.; Faisal, F.; Asif, M.A.A.R.; Hoque, M.A. Detection of Parkinson's Disease by Employing Boosting Algorithms. In *2021 Joint 10th International Conference on Informatics, Electronics and Vision, ICIEV 2021 and 2021 5th International Conference on Imaging, Vision and Pattern Recognition, icIVPR 2021*; IEEE: Piscataway, NJ, USA, 2021.
32. Demir, F.; Siddique, K.; Alswaitti, M.; Demir, K.; Sengur, A. A Simple and Effective Approach Based on a Multi-Level Feature Selection for Automated Parkinson's Disease Detection. *J. Pers. Med.* **2022**, *12*, 55. [CrossRef] [PubMed]
33. Açıcı, K.; Erdaş, Ç.B.; Aşuroğlu, T.; Toprak, M.K.; Erdem, H.; Oğul, H. A Random Forest Method to Detect Parkinson's Disease via Gait Analysis. In *Communications in Computer and Information Science*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 744, pp. 609–619.
34. Abdulhay, E.; Arunkumar, N.; Narasimhan, K.; Vellaiappan, E.; Venkatraman, V. Gait and Tremor Investigation Using Machine Learning Techniques for the Diagnosis of Parkinson Disease. *Future Gener. Comput. Syst.* **2018**, *83*, 366–373. [CrossRef]
35. Balaji, E.; Brindha, D.; Balakrishnan, R. Supervised Machine Learning Based Gait Classification System for Early Detection and Stage Classification of Parkinson's Disease. *Appl. Soft Comput.* **2020**, *94*, 106494. [CrossRef]
36. El Maachi, I.; Bilodeau, G.A.; Bouachir, W. Deep 1D-Convnet for Accurate Parkinson Disease Detection and Severity Prediction from Gait. *Expert Syst. Appl.* **2020**, *143*, 113075. [CrossRef]

37. Setiawan, F.; Lin, C.W. Implementation of a Deep Learning Algorithm Based on Vertical Ground Reaction Force Time–Frequency Features for the Detection and Severity Classification of Parkinson's Disease. *Sensors* **2021**, *21*, 5207. [CrossRef]
38. Priya, S.J.; Durai, C.A.D.; Subathra, M.S.P.; George, S.T.; Jeyabose, A. Automated Parkinson's Disease Diagnosis Using Decomposition Techniques and Deep Learning for Accurate Gait Analysis. *IEEE Access* **2025**, *13*, 74078–74091. [CrossRef]
39. Mostafavi, S.M.; Soumma, S.B.; Peterson, D.; Mehta, S.H.; Ghasemzadeh, H. Detection and Severity Assessment of Parkinson's Disease Through Analyzing Wearable Sensor Data Using Gramian Angular Fields and Deep Convolutional Neural Networks. *Sensors* **2025**, *25*, 3421. [CrossRef]
40. Trabassi, D.; Serrao, M.; Varrecchia, T.; Ranavolo, A.; Coppola, G.; De Icco, R.; Tassorelli, C.; Castiglia, S.F. Machine Learning Approach to Support the Detection of Parkinson's Disease in IMU-Based Gait Analysis. *Sensors* **2022**, *22*, 3700. [CrossRef] [PubMed]
41. Camps, J.; Samà, A.; Martín, M.; Rodríguez-Martín, D.; Pérez-López, C.; Moreno Arostegui, J.M.; Cabestany, J.; Català, A.; Alcaine, S.; Mestre, B.; et al. Deep Learning for Freezing of Gait Detection in Parkinson's Disease Patients in Their Homes Using a Waist-Worn Inertial Measurement Unit. *Knowl.-Based Syst.* **2018**, *139*, 119–131. [CrossRef]
42. Khatamino, P.; Canturk, I.; Ozyilmaz, L. A Deep Learning-CNN Based System for Medical Diagnosis: An Application on Parkinson's Disease Handwriting Drawings. In *2018 6th International Conference on Control Engineering & Information Technology (CEIT)*; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.
43. Gil-Martín, M.; Montero, J.M.; San-Segundo, R. Parkinson's Disease Detection from Drawing Movements Using Convolutional Neural Networks. *Electronics* **2019**, *8*, 907. [CrossRef]
44. Kamran, I.; Naz, S.; Razzak, I.; Imran, M. Handwriting Dynamics Assessment Using Deep Neural Network for Early Identification of Parkinson's Disease. *Future Gener. Comput. Syst.* **2021**, *117*, 234–244. [CrossRef]
45. Moshkova, A.; Samorodov, A.; Voinova, N.; Volkov, A.; Ivanova, E.; Fedotova, E. Parkinson's Disease Detection by Using Machine Learning Algorithms and Hand Movement Signal from LeapMotion Sensor. In *2020 26th Conference of Open Innovations Association (FRUCT)*; IEEE: Piscataway, NJ, USA, 2020; pp. 321–327.
46. Belić, M.; Radivojević, Z.; Bobić, V.; Kostić, V.; Đurić-Jovičić, M. Quick Computer Aided Differential Diagnostics Based on Repetitive Finger Tapping in Parkinson's Disease and Atypical Parkinsonisms. *Heliyon* **2023**, *9*, e14824. [CrossRef]
47. Ul Haq, A.; Li, J.; Memon, M.H.; Khan, J.; Ali, Z.; Abbas, S.Z.; Nazir, S. Recognition of the Parkinson's Disease Using a Hybrid Feature Selection Approach. *J. Intell. Fuzzy Syst.* **2020**, *39*, 1319–1339. [CrossRef]
48. Sabeena, B.; Sivakumari, S.; Teressa, D.M. Optimization-Based Ensemble Feature Selection Algorithm and Deep Learning Classifier for Parkinson's Disease. *J. Healthc. Eng.* **2022**, *2022*, 1487212. [CrossRef]
49. Silva, G.; Rodrigues, P.M.; Alves, M.; Cunha, R.; Bispo, B.C.; Rodrigues, P.M. Parkinson Disease Early Detection Using EEG Channels Cross-Correlation. *Int. J. Appl. Eng. Res.* **2020**, *15*, 197–203.
50. Al-Jabbar, M.; Alshahrani, M.; Senan, E.M.; Abunadi, I.; Almalki, S.A.; Alshari, E.A. Hybrid Techniques of Multi-CNN and Ensemble Learning to Analyze Handwritten Spiral and Wave Drawing for Diagnosing Parkinson's Disease. *Comput. Model. Eng. Sci.* **2025**, *143*, 2429–2457. [CrossRef]
51. Bahaddad, A.A.; Ragab, M.; Ashary, E.B.; Khalil, E.M. Metaheuristics with Deep Learning-Enabled Parkinson's Disease Diagnosis and Classification Model. *J. Healthc. Eng.* **2022**, *2022*, 9276579. [CrossRef]
52. Cincovic, J.; Jovanovic, L.; Nikolic, B.; Bacanin, N. Neurodegenerative Condition Detection Using Modified Metaheuristic for Attention Based Recurrent Neural Networks and Extreme Gradient Boosting Tuning. *IEEE Access* **2024**, *12*, 26719–26734. [CrossRef]
53. Markovic, F.; Jovanovic, L.; Spalevic, P.; Kaljevic, J.; Zivkovic, M.; Simic, V.; Shaker, H.; Bacanin, N. Parkinsons Detection from Gait Time Series Classification Using Modified Metaheuristic Optimized Long Short Term Memory. *Neural Process. Lett.* **2025**, *57*, 14. [CrossRef]
54. Arora, S.; Venkataraman, V.; Zhan, A.; Donohue, S.; Biglan, K.M.; Dorsey, E.R.; Little, M.A. Detecting and Monitoring the Symptoms of Parkinson's Disease Using Smartphones: A Pilot Study. *Park. Relat. Disord.* **2015**, *21*, 650–653. [CrossRef] [PubMed]
55. Sharma, V.; Mankodiya, K.; De, F.; Torre, L.; Zhang, A.; Ryan, N.; Ton, T.G.N.; Gandhi, R.; Jain, S. *SPARK: Personalized Parkinson Disease Interventions Through Synergy Between a Smartphone and a Smartwatch*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2014; Volume 8519.
56. Belic, M.; Bobic, V.; Badza, M.; Solaja, N.; Djuric-Jovicic, M.; Kostic, V.S. Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—A review. *Clin. Neurol. Neurosurg.* **2019**, *184*, 105442. [CrossRef] [PubMed]
57. Mei, J.; Desrosiers, C.; Frasnelli, J. Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature. *Front. Aging Neurosci.* **2021**, *13*, 633752. [CrossRef]
58. Zadoo, S.; Singh, Y.; Singh, P. Automated Parkinson's Disease Detection: A Review of Techniques, Datasets, Modalities, and Open Challenges. *Int. J. Smart Sens. Intell. Syst.* **2024**, *17*, 1–37. [CrossRef]

59. Rabie, H.; Akhloufi, M.A. A review of machine learning and deep learning for Parkinson's disease detection. *Discov. Artif. Intell.* **2025**, *5*, 24. [CrossRef]
60. Patro, P.; Das, T. Artificial Intelligence in Action (2019–2024): A Review on Parkinson's Disease Detection Using Non-Invasive Procedures. *IEEE Access* **2025**, *13*, 116586–116605. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.