




Review

# Survey of Machine Learning Techniques in the Analysis of EEG Signals for Parkinson's Disease: A Systematic Review

Ana M. Maitin <sup>1</sup> , Juan Pablo Romero Muñoz <sup>2,3</sup>  and Álvaro José García-Tejedor <sup>1,\*</sup> 

<sup>1</sup> Centro de Innovación Experimental del Conocimiento (CEIEC), Universidad Francisco de Vitoria, 28223 Madrid, Spain; a.maitin@ceiec.es

<sup>2</sup> Facultad de Ciencias Experimentales, Universidad Francisco de Vitoria, 28223 Madrid, Spain; p.romero.prof@ufv.es

<sup>3</sup> Brain Damage Unit, Hospital Beata María Ana, 28007 Madrid, Spain

\* Correspondence: a.gtejedor@ceiec.es

**Abstract:** Background: Parkinson's disease (PD) affects 7–10 million people worldwide. Its diagnosis is clinical and can be supported by image-based tests, which are expensive and not always accessible. Electroencephalograms (EEG) are non-invasive, widely accessible, low-cost tests. However, the signals obtained are difficult to analyze visually, so advanced techniques, such as Machine Learning (ML), need to be used. In this article, we review those studies that consider ML techniques to study the EEG of patients with PD. Methods: The review process was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, which are used to provide quality standards for the objective evaluation of various studies. All publications before February 2022 were included, and their main characteristics and results were evaluated and documented through three key points associated with the development of ML techniques: dataset quality, data preprocessing, and model evaluation. Results: 59 studies were included. The predominating models were Support Vector Machine (SVM) and Artificial Neural Networks (ANNs). In total, 31 articles diagnosed PD with a mean accuracy of  $97.35 \pm 3.46\%$ . There was no standard cleaning protocol for EEG and a great heterogeneity in EEG characteristics was shown, although spectral features predominated by 88.37%. Conclusions: Neither the cleaning protocol nor the number of EEG channels influenced the classification results. A baseline value was provided for the PD diagnostic problem, although recent studies focus on the identification of cognitive impairment.

**Keywords:** machine learning; deep learning; artificial neural networks; electroencephalogram; Parkinson's disease; review



**Citation:** Maitin, A.M.; Romero Muñoz, J.P.; García-Tejedor, Á.J. Survey of Machine Learning Techniques in the Analysis of EEG Signals for Parkinson's Disease: A Systematic Review. *Appl. Sci.* **2022**, *12*, 6967. <https://doi.org/10.3390/app12146967>

Academic Editor: João M. F. Rodrigues

Received: 10 June 2022

Accepted: 7 July 2022

Published: 9 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Parkinson's disease (PD) is one of the most common chronic progressive neurological disorders, affecting between 7 million and 10 million people worldwide [1]. It is characterized by the loss of dopaminergic neurons in the *substantia nigra* [2], and it is not until advanced stages of neurodegeneration (when the patient has a 50–70% neuronal loss in the *substantia nigra* [3,4]) that the characteristic motor symptoms of this disease (bradykinesia, rigidity, and tremor at rest [5]) appear. Consequently, the first line of treatment includes the administration of Levodopa [6,7], a drug that compensates for the loss of dopamine. Regarding the diagnosis of PD, it is clinical and is based on motor symptoms, and thus is carried out through clinical assessment of repetitive limb movements, resistance to passive mobilization, spontaneous movements, balance, and gait pattern. The diagnosis requires the evaluation of an experienced clinician and a high degree of suspiciousness for the disease, and misdiagnosis in the early stages is not rare [8]. Hence, advances in new diagnostic techniques in PD could help to detect this disease in earlier stages, allowing for anticipating the administration of dopaminergic medication, which may favor the quality of life of patients during a longer period.

There is a wide variety of techniques in the field of neurology that could be used to support Parkinson's clinical diagnosis. These include image-based tests such as SPECT and Cardiac MIBG, both revealing indirect physiological consequences of dopaminergic denervation. EEG tests are non-invasive techniques, low-cost compared to the mentioned image-based tests, and they are present in most health centers. This technique records the electrical activity of pyramidal neurons of the brain cortex and has shown to be effective in other neurological diseases such as in the diagnosis or prediction of epileptic seizures, in the development of biomarkers of Alzheimer's disease, or even experimentally in the detection of abnormalities in schizophrenia [9–11]. Moreover, it has a high temporal resolution and a high test-retest, that is, it is capable of reproducing results regardless of external aspects.

However, two main characteristics define the EEG signals and make their subsequent analysis difficult: the low signal-to-noise ratio and their stochastic nature. A low signal-to-noise ratio indicates a high level of noise in EEG signals, so pre-processing is required to filter out the signal noise and remove possible artifacts and then analyze the signals and obtain results. The main problem when filtering the signals is that there is no standard protocol to perform the cleaning. Another problem comes from the fact that when the signal noise is removed, relevant components in EEG signals can also be eliminated, which may lead to misdiagnosis. Regarding the stochastic character, it indicates that the state of occurrence of an event does not depend on the previous event. Hence, to extract the essential characteristics of the signals, advanced techniques must be used for the study of nonlinear dynamics, which require more expensive computational methods [12,13]. Artificial Intelligence (AI) models have shown to be some of the most suitable methods to deal with these difficulties since their increasing development in recent years has made them indispensable tools for analyzing and understanding the large amount of data generated by today's society.

As a discipline, AI encompasses many techniques but it is the field of Machine Learning (ML) that currently provides the most promising results. ML is a scientific discipline that studies and develops algorithms capable of generalizing behaviors and recognizing hidden patterns in a large amount of data by way of examples. It was defined by Arthur Samuel as: "field of study that gives computers the ability to learn without being explicitly programmed" [14]. ML techniques can be classified into two categories depending on the type of processing that is carried out: symbolic processing, which uses formal languages, logical orders, and symbols, and subsymbolic processing, which is designed to estimate functional relationships between data. These techniques are receiving increasing interest from the medical domain, where they have been mostly used in image analysis [15], although in recent years their application has spread to other areas [16,17].

Within ML techniques, Deep Learning (DL) has been a breakthrough in the last years. DL is performed by Deep Neural Networks, a subset of Artificial Neural Networks (ANN) that are biologically inspired by how human neurons work. It is defined in [18] as multiple-layer, hierarchical models that can learn representations of data with multiple levels of abstraction. ANN (and so DL networks) can learn from data using three different strategies: supervised, unsupervised, or reinforced. However, all three require a large amount of input data and a careful training process to learn, as ANNs are well-known to be as powerful as their training data. Therefore, proper selection of the dataset is paramount as well as input data transformation using mathematical operations. So, input dataset quality has also been checked in the reviewed articles.

This review analyzes the current impact of the use of ML techniques for the study of EEGs of patients with PD, with the aim of serving as a starting point for researchers in future studies related to this disease, providing reference values obtained from a comprehensive quality criterion, and discovering the current trend of this research topic so that it favors the development of new applications in the clinical setting. Specifically, the process of selection of the reviewed articles has followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [19], a methodology based on a search in different databases, and the application of significant exclusive criteria from the fields of medicine

and computational sciences. Although numerous articles apply ML and DL techniques to study different diseases through EEG analysis, such as epilepsy, schizophrenia, Alzheimer's, and sleep EEG [20–27], its use associated with Parkinson's disease is not so widespread, and it has not been until recently that the literature in this field has begun to be developed. In a previous work [28] we reviewed, from a medical point of view, the literature in this field with these same search terms, focuses the analysis on the clinical state of patients, and the recording parameters, protocol, and test of the EEG, without delving into the computational techniques used. Moreover, such work was restricted to EEG acquired in the resting state and the motor activation tests. On the contrary, we widen the study selection without restrictions on their objectives or the EEG tests performed, with a special emphasis on computational development. Hence, it provides a summary of the current state of use of ML methods in the EEGs of patients with PD, as it includes the analysis of the techniques considered by the studies, their methodology, architectures developed and the results obtained. The conclusions derived from this analysis may serve as an entry point and a reference for future work on EEG markers search for Parkinson's Disease.

This review is divided into the following sections: Methods, which includes the methodology used, articles found, databases used, and the inclusion and exclusion criteria that were applied to select the articles analyzed in this review; Results, which extracts, from the selected articles, the types of tests carried out, number of patients, EEG cleaning protocol, feature extraction, ML techniques used, validation methods, and results of each model; and Discussion, which compares the information in the selected articles to draw the conclusions of the review.

## 2. Methods

### 2.1. Search Strategy and PRISMA Methodology

The review process performed in this article followed the PRISMA guidelines, defined in [19] as “an evidence-based minimum set of items aimed at helping authors to report a wide array of systematic reviews and meta-analyses that assess the benefits and harms of a health care intervention”. This methodology provides quality standards through items related to the title, the abstract, the methods, the results, and the discussion, through which the authors can make critical and objective evaluations of various studies.

The selection process was carried out in different phases. The first step consisted in determining the keywords to perform the searches in the databases so that they fitted the topic proposed for the review based on both medical and computational points of view. From the medical point of view, the search topics were Parkinson's disease and electroencephalography. The specific terms chosen were: 1. Parkinson, 2. EEG, and 3. electroencephalogram. Regarding computational terms, the analyzed publications should use tools and techniques from the field of Artificial Intelligence. ML terminology does not always appear within the articles in the same way, so some of the most widely used generic terms were considered to cover as many works as possible. These were: 4. machine learning, 5. deep learning, and 6. neural networks. The proposed search terms were combined using logical operators as follows: 1 AND (4 OR 5 OR 6) AND (2 OR 3). This combination was introduced in the following 5 databases: Web of Science, PUBMED, Scopus, CINAHL, and Science Direct. The search was performed on 14 February 2022, with no time limit, providing a total of 358 results, which were downloaded to the Zotero platform for further analysis.

The second step consisted of the screening phase, in which, after the removal of duplicate elements, contents based on academic books, book chapters, abstracts, and posters were discarded as they were considered to be outside the scope of this review. Next, in the eligibility phase, the exclusion criteria shown in Table 1 were applied, and ordered according to the objectives proposed for this review. The reasons for which these criteria were chosen are also included in Table 1. To determine whether or not a study satisfied a criterion, its title, its abstract, and the full article, if necessary, were analyzed. Even though all studies contained the specified search terms, some of the studies only used them as a

reference to other works, or mention them without developing them. These cases were classified as studies that did not use ML techniques, studies that did not consider EEG, or studies that did not focus on PD.

**Table 1.** Exclusion criteria and reason for the exclusion.

Exclusion Criteria	Reason for Exclusion
Studies focused on neurological diseases different from PD	Outside the objectives of this review
Studies not using ML techniques	Outside the objectives of this review
Studies not considering EEG	Outside the objectives of this review
Studies with invasive EEG	Invasive EEG signals are not comparable with non-invasive EEG ones. Moreover, non-invasive EEG was chosen for being low-cost, widely available, and easy to acquire, which are properties not shared by invasive EEG.
Studies on animals	The results obtained in studies with animals may not be always extrapolated to humans. Moreover, EEG of animals and humans are not comparable.
Pharmacological studies	The studies focused on the development and analysis of the components of medications are outside the objectives of this review.
Review articles	Outside the objectives of this review

After the eligibility process, the inclusion phase was the last step of the selection procedure. This phase received, on the one hand, the studies that passed the exclusion criteria, and on the other hand, the studies that were not found or were retracted and thus could not go through the eligibility step. This last group of studies was discarded and, as a result, the inclusion phase provided us with the articles that constituted this review and that were considered for further analysis.

## 2.2. Data Extraction and Analysis

After the previous selection process, for each selected article, the information associated with the following topics was extracted in Table 2:

**Table 2.** Description of the items considered for the analysis of the articles.

Item	Description
1. Dataset quality	Through clinical and technical parameters such as the number of patients in the study, the stage of the disease, the administration of medication, and the type of EEG tests performed
2. Pre-processing of data	Through the EEG cleaning protocol and feature extraction methods
3. Analysis of ML techniques	Through the types of models, model architecture, evaluation of the quality of the training/validation process, metrics used, and results of each model

These topics were chosen to synthesize the most relevant information within each of the articles, to analyze each item, and provide a starting point for future studies.

The conclusions derived in this review were obtained, on the one hand, by comparing, for each of these points, the information collected in the different articles, and, on the other hand, by evaluating the results obtained by each article about the parameters used. To perform this analysis, the Matplotlib (<https://matplotlib.org> (accessed on

14 February 2022)) library in Python has been used to make the figures, and the Numpy (<https://numpy.org> (accessed on 14 February 2022)) and Scipy (<https://scipy.org> (accessed on 14 February 2022)) libraries in Python have been used to develop the algorithms that provide the statistical results.

### 3. Results

#### 3.1. PRISMA Flow Diagram

As shown in the PRISMA flow diagram displayed in Figure 1, the search process in the databases yielded 358 results, 132 of which were duplicates, and thus, were eliminated. The remaining 226 results were then submitted to the screening process, where 25 academic books, one abstract, and one poster, were rejected. As a consequence, 27 studies were removed, leaving a total of 199 articles for the eligibility phase, where the exclusion criteria described in Table 1 were applied. To implement this process, the title and abstract of the research articles were reviewed first, and, if doubts about the exclusion criteria arose, a complete read was carried out. As a result of this phase, 24 articles were found to not use ML techniques, 27 articles did not focus their study on PD, 29 articles did not use EEG techniques, two articles considered animal studies, seven articles were pharmacological, 43 articles were reviews with a different purpose, and six studies considered invasive EEG. As detailed in the PRISMA diagram in Figure 1, the sum of all these types of excluded articles resulted in a total of 138 rejections, leaving 61 articles for the inclusion phase. Finally, one article obtained from the previous process was retracted, and one article could not be found, so both were excluded from the final result. Therefore, 59 articles were finally obtained for their subsequent analysis.

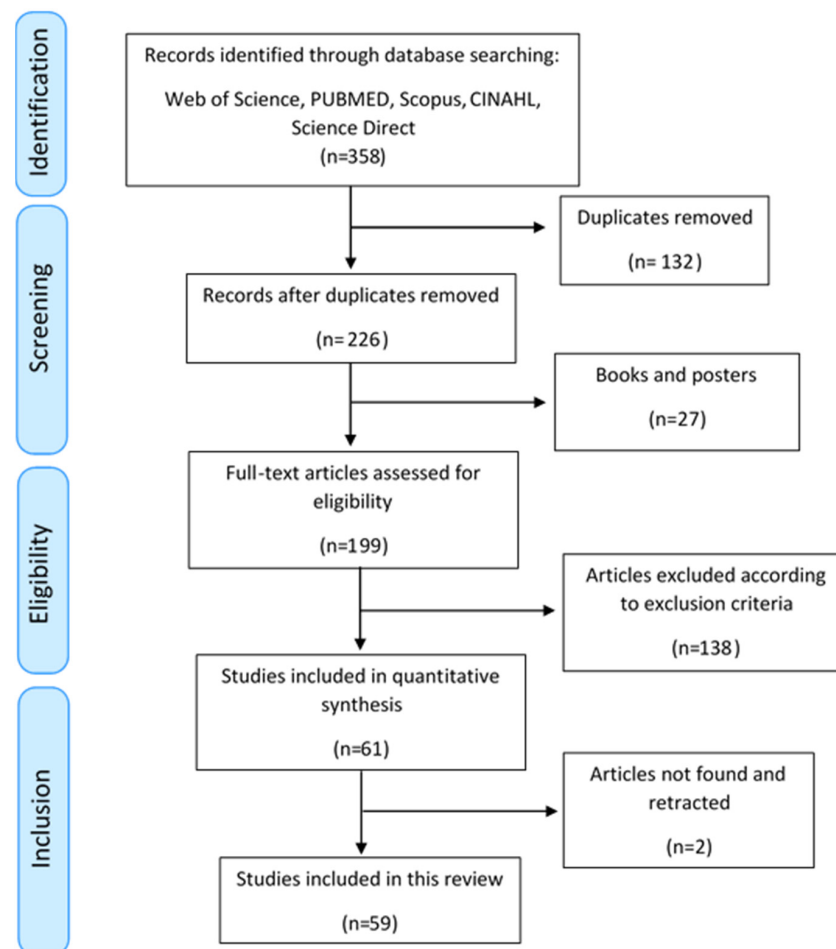
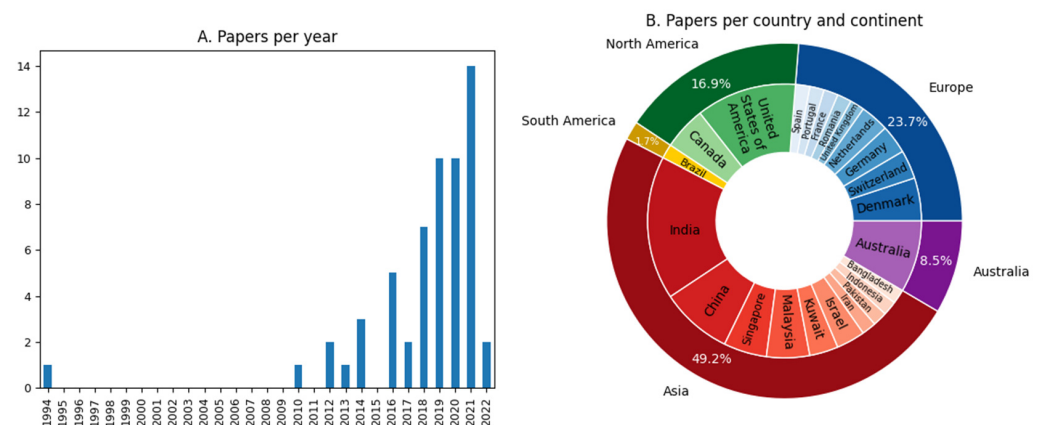


Figure 1. PRISMA diagram of the bibliographic review conducted.

### 3.2. Statistical Analysis

ML techniques have been a growing development in recent years, increasing their use in different areas. This growth, particularized in the application of ML techniques on EEG associated with PD, was captured by the articles selected in this review, as displayed in Figure 2A, which shows an increase in publications, and so a growing interest in this topic in the last 5 years, whereas in previous years, the development was intermittent. A decrease in the production of the number of studies in 2022 compared to the previous year can be appreciated, although it must be taken into account that 2022 was not fully included in the review due to the search date (14 February). So 2022 results are not comparable with previous years. It should also be remarked that in Figure 2A, the last publication date has been considered.



**Figure 2.** (A) Bar plot of number of selected articles per year taking into account their issue publication. (B) Pie-chart with the distribution of the selected articles according to the country and continent associated with the first affiliation of the first author.

Figure 2B shows the distribution of the articles according to the (first) affiliation country of the first author, as well as the continent to which such country belongs. The continent that contributed the most to the development of new investigations of PD by means of EEG using ML techniques was Asia (49.2%), followed by Europe (23.7%). Although the contributions from North America (16.9%) and Australia (8.5%) were smaller in proportion, it is worth emphasizing the fact that there was global scientific interest in this topic in the last few years. Regarding the country, it can be appreciated that both the USA and India contributed the largest number of publications within their respective continent in similar proportions. Both China and Australia had a similar contribution, whereas in Europe, the distribution was homogeneous.

Delving into the content of the articles selected for this review, it was necessary to fix a common framework to extract the most relevant information. In Table 3, this information was summarized according to three key points chosen by their importance within the area of ML. These were:



**Table 3.** Summary of the objectives, participants, state of PD, EEG pre-processing, features, models used, model parameters, training and validation methods, and best results for each article included in this review.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[29]	Classification of PD patients vs. controls.	Subjects: 20 PD and 21 controls Age: PD: 67.6 ± 7.0 HC: 67.5 ± 6.4	HY scale: 1–2 UPDRS: 23.5 ± 9.8 Disease Duration: 7.6 ± 4.3 Medication: ON and OFF	Sixty-four-channel EEG recorded during 1 min in resting state at 1 kHz. In total, 27 electrodes were considered. Impedances were kept below 15 kΩ. The EEG was divided into non-overlapping 3 s segments.	The segments of EEG were introduced as input.	CNN + RNN	Two 1D-conv layers with 64 each, LSTM with 80 cells and a fully connected layer with 50 units. The activation function was sigmoid, the loss function was “binary crossentropy” and the optimizer was Adam with learning rate = 0.001.	Data were split into 80% for training and 20% for test sets.	Results without medication: Accuracy = 96.9 Precision = 100 Recall = 93.4
[30]	Classification of PD patients vs. controls.	Subjects: 20 PD and 21 controls Age: PD: 67.6 ± 7.0 HC: 67.5 ± 6.4	HY scale: 1–2 UPDRS: 23.5 ± 9.8 Disease Duration: 7.6 ± 4.3 Medication: ON and OFF	Sixty-four-channel EEG recorded during 1 min in resting state at 1 kHz. In total, 27 electrodes were considered. Impedances were kept below 15 kΩ. A band-pass filtered was used at 1–55 Hz, and re-referenced to average reference. Artifacts were removed using ICA. The data were standardized and segmented into 1 s or 2 s epochs.	For the CNN/CRNN models, 1 s segments without overlapping were introduced. For the others, two datasets were considered. First, with 2 s epochs and 13 features of HOS. Second, 794 time-series features in 1 s epochs using the bands theta, alpha beta, and gamma. Top significant features were selected through ANOVA.	CNN + RNN, CNN, KNN, SVM, RF	The final architecture for KNN, SVM, and RF were not specified. CNN + RNN: 2 1D-conv (kernel size 3, filters 32 and 64), max-pooling, a GRU cells with 35 units, and 2 fully connected layers (time distributed and dense with 35 units). Dropout were used. Activation function was ReLU and a softmax for the final layer. Optimizer Adam (learning rate = 0.001) and a binary cross-entropy as loss. CNN: 4 consecutive blocks, with 8, 12, 12, and 16 filters in each 1D-conv layer (kernel size = 9). Max-pooling layer. Three fully connected layers with 30 and 5 units, activation function ReLU, and softmax for the last one. Adam optimizer (learning rate = 0.001), and binary cross-entropy as loss function.	Nested cross-validation (inner 5-fold for hyperparameter tuning and outer 10 fold)	The CNN + RNN model obtained: Accuracy = 99.2 Precision = 98.9 Recall = 99.4 F1-score = 99.2 AUC = 99.2
[31]	Classification of PD patients vs. controls.	Subjects: 20 PD and 20 controls. Age: PD: (45–65) HC: 58.1 ± 2.95	HY scale: 1: n = 2; 2: n = 11; 3: n = 7 UPDRS: Not specified Disease Duration: 5.75 ± 3.52 Medication: ON	Fourteen-channel EEG recorded during 5 min in resting state at 128 Hz. Epochs of 2 s were segmented and a threshold technique was applied at ±100 μV. A band-pass filter was used at 1–49 Hz.	The EEG signals were the input.	CNN	Thirteen layers with 4 1D-conv layers, 4 max-pooling layers, and 3 fully connected layers. Adam optimizer (learning rate = 10 <sup>-4</sup> ). Activation function Relu and the last one softmax. Dropout of 0.5.	Ten-fold cross-validation with stratified data. In total, 20% of the training data was also used for validation at the end of each epoch.	The results were: Accuracy = 88.25 Sensitivity = 84.71 Specificity = 91.77

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[32]	Classification of PD patients with MCI vs. patients without MCI.	Subjects: 27 PD with MCI and 43 PD without MCI Age: MCI: (53–84) Non-MCI: (46–82)	HY scale: MCI: 0–5 Non-MCI: 0–4 UPDRS: MCI: 0–41 Non-MCI: 0–36 Disease Duration: MCI: 0–23 Non-MCI: 0–17 Medication: Not specified	Two-hundre-fifty-six-channel EEG recorded during 12 min in resting state with eyes-closed at 1 kHz. In total, 214 electrodes were associated with 10 ROI. Signals were filtered at 0.5–70 Hz, with a 50 Hz notch, and an inverse Hanning window was used to stitch together segments to get 3 min of EEG data. Artifacts were removed. The average of all “good” channels was used to reference the signals to an average value.	The spectral power was calculated in 10 ROI and globally for 6 frequency intervals resulting in 66 spectral features. The PLI was calculated between all pairs of ROI resulting in 330 connectivity features.	RF	The standard implementation in R was applied.	Data were split into 70% for training and 30% for test sets. 20 runs of 5-fold cross-validation.	The results with the combination of both features. ROI Train AUC = 0.73 ± 0.16 Test AUC = 0.71 Without ROI: Train AUC = 0.7 ± 0.14 Test AUC = 0.875
[33]	Selection of the QEEG parameters that best distinguish between controls and PD patients.	Subjects: 50 PD and 41 controls. Age: PD: 68.8 ± 7 HC: 71 ± 7	HY scale: Not specified UPDRS: Not specified Disease Duration: 5.3 ± 5.1 Medication: Not specified	Two-hundred-fifty-six-channel EEG recorded during 12 min in resting state with eyes-closed at 500 Hz. Three min were constructed with segments of at least 30 s without artifacts, and a 0.5–70 Hz filter was applied. An inverse Hanning window was used to join segments. It was referenced with respect to mean and defective channels that were interpolated with the spherical spline method. Artifacts were removed.	Ten brain regions were considered with 79 different measurements. All features were extracted from the frequency spectrum.	RF, SVM, DT, LR, LR with LASSO	SVM: RBF kernel. Optimization was carried out for tuning parameters.	Ten-fold cross-validation.	The most significant models were: RF: Accuracy = 78 AUC = 0.8 LR with LASSO: AUC = 0.76
[34]	Cognition classification of patients with PD	Subjects: 20 PD H-COG, 20 PD L-COG and 72 inter-COG Age: H-COG: 59.5 (54.6–66.4) L-COG: 67.8 (60.1–72.1) Inter-COG: 63.5 (57.7–68.0)	HY scale: Not specified UPDRS: H-COG: 18.5 L-COG: 23 Inter-COG: 20.5 Disease Duration: H-COG: 11.2 ± 4.5, L-COG: 10.9 ± 5.1, Inter-COG: 11.8 ± 8.0 Medication: ON	Twenty-one-channel EEG recorded in resting state with eyes-closed. Data were re-referenced. After visual confirmation of artefact-free signals, 5 consecutive non-overlapping 4096-point (8.192 s) epochs were selected. Recordings with less than five epochs were excluded.	In total, 16,674 features were extracted per patient. Feature-selection was performed using a Boruta algorithm. Small feature sets were considered as input.	RF	The hyperparameters were optimized with a variant of Bayesian Optimization technique called Mixed Integer Parallel Efficient Global Optimization (MIP-EGO) for mixed-integer categorical search spaces.	Ten-fold cross-validation. Additional assessment with a combination of cross-validation and split-sample validation.	Using all features from all cross-validation runs in L-COG vs. H-COG: Accuracy: 92 Sensitivity: 90 Specificity: 94



Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[35]	Prediction of FOG episodes	Subjects: 16 PD Age: PD: $70.88 \pm 6.92$	HY scale: $2.75 \pm 0.61$ UPDRS: $42.50 \pm 14.25$ Disease Duration: $8.63 \pm 6.58$ Medication: Not specified	Four-channel EEG during 404 FOG episodes in structured series of Timed Up and Go tasks at 500 Hz, with a duration between 1 and 220 s. Segments with artefacts were rejected using visual inspection. In total, 1902 selected samples of data were filtered using band-pass (0.5–60 Hz) and band-stop (50 Hz) Butterworth IIR filters with zero phase shift. The data were normalized and ICA was applied.	DTF was applied. The non-parametric Wilcoxon Sum Rank Test was used to select the most significant feature. A $p$ -value < 0.05 and $r$ -value > 0.25 were chosen for further processing.	BNN	Three layers Back Propagation BNN was used as a classifier with Bayesian regularisation and Levenberg-Marquardt optimization.	Eleven patients were randomly chosen. Fifty runs of random training/validation (50%) and test (50%). Remaining 5 patients were considered for test.	In train set: Sensitivity: 82.65 Specificity: 86.60 In test set: Sensitivity: 85.86 Specificity: 80.25
[36]	Classification of early-stage PD patients vs. HC	Subjects: 19 PD and 30 HC Age: PD: $63.7 \pm 7.8$ HC: $64.4 \pm 6.2$	HY scale: $1.8 \pm 0.6$ UPDRS: $20.1 \pm 8.8$ Disease Duration: $1.1 \pm 0.9$ Medication: OFF	Sixty-four-channel EEG recorded at 250 Hz while performing visual Go/No-Go and AOB during 15 min in cognitive tasks. Signals were referenced to average mastoid electrodes and band-pass filtered in 4 bands with overlap. Then they were cut into epochs based on stimulus onset and response, and averaged across trials of the same condition.	In total, 199 features extracted by the BNA analysis from the HC and PD groups were used. In each iteration, the FPR feature selection method was applied.	LR	Not specified	Ten-fold cross-validation with stratified data.	Cross validation results in discriminating HC vs. early stage PD: AUC: 79 Sensitivity: 74 Specificity: 73
[37]	Classification of PD patients vs. controls using 6 emotional stimuli.	Subjects: 20 PD and 20 controls Age: PD: (40–65) HC: (40–65)	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Fourteen-channel EEG recorded during 6 emotional stimuli. The signals were segmented into 10 s epochs with overlapping of 75%. Then, pass-band elliptic filters were used to obtain the alpha, beta, and gamma bands.	Three spectral features were calculated, that is, Spectral Entropy (SEN), Spectral Energy-Entropy (SEEN), and SpectralTeager Energy-Entropy (STEEN) for each band.	PNN, KNN, SVM	KNN: $k$ between 1 and 10. PNN: A multilayer feed-forward network with 4 layers was considered. It used an exponential activation function with $\sigma$ ranged from 0.55 to 0.65.	Not specified	The best accuracy for each emotion: Happiness: 96.8 with PNN and SEEN Sadness: 90.2 with KNN and SEN Fear: 95.07 with KNN and SEEN Anger: 91.4 with SVM and SEEN Surprise: 94.53 with KNN and SEEN Disgust: 88.18 with SVM and SEEN.

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[38]	Classification of PD patients vs. controls	Subjects: Dataset1: 15 PD and 16 HC Dataset2: 20 PD and 20 HC Age: Dataset1: PD: $63.5 \pm 9.6$ HC: $63.2 \pm 8.2$ Dataset2: PD: $58.1 \pm 2.95$ HC: $59.05 \pm 5.94$	HY scale: Dataset1: 2 and 3 Dataset2: 1–3 UPDRS: Not specified Disease Duration: Dataset1: $4.5 \pm 3.5$ Dataset2: $5.75 \pm 3.52$ Medication: Dataset1: ON and OFF	Dataset1: 32-channel EEG recorded during 3 min in resting state at 512 Hz. Artifacts were manually removed and a highpass filter at 0.5 Hz was used. Dataset2: 14-channel EEG during 5 min in resting state at 128 Hz. Signals were segmented into 2 s windows. Eye blinking artifacts were removed with a threshold at $\pm 100$ V and a forward and reverse filtering technique using sixth-order Butterworth filter at 1–49 Hz was used.	EEG signals are subjected to SPWVD to obtain TFR. After Kaiser window selection and resizing, the two-dimensional plots are fed to the model. Experiments are carried out by maintaining the same setup for both datasets.	CNN	Four 2D-Conv Layers, 2 MaxPooling layers, 2 Fully Connected Layers (50 and 32 neurons), and a Softmax layer. The number of filters selected were 96, 32, 16, and 8. A filter size of $7 \times 7$ , $5 \times 5$ , and $3 \times 3$ with a stride of 2 was used. A dropout of 0.5 was considered. Adam optimizer was used with learning rate = $10^{-4}$ .	Ten-fold cross validation	The best results were for dataset1 HC vs. PD ON medication: Accuracy: 100 Specificity: 100 Sensitivity: 100 Precision: 100 F1 Score: 100
[39]	Classification of controls vs. PD patients with ON and OFF medication	Subjects: 15 PD and 16 HC Age: PD: $63.2 \pm 8.2$ HC: $63.5 \pm 9.6$	HY scale: 2 or 3 UPDRS: ON: $33.7 \pm 10.9$ OFF: $45.5 \pm 13.0$ Disease Duration: Not specified Medication ON and OFF	Thirty-two-channel EEG during 3 min focusing on an image at 512 H.	EEG recordings were split in half and then converted into spectrograms using Gabor transform.	CNN	Two-dimensional-Conv (16 filters kernel $5 \times 5$ , ReLu), Dropout (0.2), 2D-Conv (32 filters, kernel $3 \times 3$ , ReLu), MaxPooling, Flatten, Dense (unit size of 512, ReLu), Dropout (0.7), Dense (unit size of 3 for Softmax and 1 for sigmoid). Adam optimizer was used with a learning rate = 0.001 and a decay rate of 0.01.	10-fold cross validation	Results for 3 class classification: Accuracy: $99.46 \pm 0.73$ Precision: $99.48 \pm 0.01$ Sensitivity: $99.46 \pm 0.01$ F1 Score: $99.46 \pm 0.01$
[40]	Determine the optimal montage to detect FOG.	Subjects: 7 PD Age: Not specified	HY scale: Not specified UPDRS: Not specified DiseaseDuration: Not specified Medication: OFF	Thirty-two-channel EEG in main cortical regions at 512 Hz during a structured series of Timed Up and Go tasks. Average of 2 ear lobes electrodes was taken as reference. Data were segmented into 1 s windows and filtered at 0.5–50 Hz. 343 s of EW and 343 s of FOG samples were collected.	Division in bands was implemented. Z-transformation was applied to normalize EEG data. Power spectral density, centroid frequency and power spectral entropy were extracted.	Feed-forward neural network	Six hidden nodes. Levenberg Marquardt's algorithm with early stopping was used.	Fifty times runs. Data was divided into training 34%, validation 33%, and testing 33%.	Results with only 2 channels C4-O2: Sensitivity: 72.54 Accuracy: 69.71
[41]	Classification of PD patients vs. controls.	Subjects: 100 PD and 100 controls. Age: PD: (50–70) HC: (50–70)	HY scale: 1–1.5 UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Two-channel EEG recorded for 30 min for the flexion and extension of the wrist. 5–50 Hz band-pass filter was applied.	EEG: Lyapunov and inverse Lyapunov exponent, Shannon Entropy EMG: power, standard deviation, root mean square, variance, waveform length, modified median, and mean frequency.	MLP	Three algorithms were tested. 1. Gradient Descent algorithms (traingd, traingdm), 2. Conjugate Gradient algorithms (trainnscg, trainingcp), and 3. Quasi-Newton algorithms (trainbfg, trainlm). Sigmoid function was used in the hidden layer. The number of hidden neurons was checked for 5, 7, 9, 10, 20, and 30.	The dataset was divided into training 70%, validation 15%, and testing 15%.	ANN with trainlm and 10 neurons: Accuracy = 100 RMSE = $4.03 \times 10^{-3}$ R value = 0.9998

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[42]	Classification of PD patients vs. controls.	Subjects: 40 PD and 30 controls Age: PD: 63.53 ± 4.95 HC: 64.72 ± 5.74	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Sixty-four-channel EEG recorded during 40 vocalizations of 5–6 s of the vowel /u/ with 5 pitch shifts each. The impedances were kept below 50 kΩ. A band-pass filter at 1–20 Hz was applied. The signals were segmented in epochs of 700 ms that contained the pitch shift. Referenced to the average of the mastoid electrodes. Trials with artifacts were rejected.	The epochs were the input data.	CNN, RNN, 2D-CNN-RNN, 3D-CNN-RNN	CNN: 8 layers. 2 -Conv, max-pooling, Conv, max-pooling, Conv, 2 fully connected layers with 1000 and 500 units. RNN: GRU layer with 6 units and 2 fully connected layers with 1280 and 300 units. 2D-CNN-RNN: Conv, GRU with 6 units and 2 fully connected layers with 2000 and 300 units. 3D-CNN-RNN: 2 Conv, max-pooling, GRU with 6 units and 2 fully connected layers with 2000 and 300 units. All models shared Adam optimizer with learning rate 0.001, ReLU activation, and softmax for the output layer. Filters, strides, and depths were specified.	Five-fold cross-validation. The trials from the same patient were involved in either training or test.	The best model was 3D-CNN-RNN: Accuracy = 82.89 ± 9.60
[43]	Classification of PD patients vs. controls.	Subjects: 27 PD and 30 controls Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Two-hundred-fifty-six-channel EEG recorded during 2 types of visual stimuli (Stim and No-Stim) of 2.4 s. 185 electrodes were selected removing the channels located on the face and neck. Cleaning was not specified.	For each trial, the FFT of the last 2.24 s was computed. The features consisted of the spectral amplitudes. Six channels over the occipital area were selected providing a total of 162 patient data and 180 control data.	LR, DT, RF	DT: Gini index measure of entropy was used, maximum depth of 5, minimum of 2 samples for split, and a minimum number of 1 sample per leaf as stopping criterion. RF: 100 trees of a max depth of 30.	For each stimulus 1000 runs of random train (70%) and validation (30%). The models were tested using the other stimulus.	For both methodologies, the best model was RF. Stim for train: AUC_val = 0.994 AUC_test = 0.71 No-Stim for train: AUC_val = 0.998 AUC_test = 0.66
[44]	Classification of PD patients vs. controls.	Subjects: 21 PD and 25 controls. Age: PD: 62.7 ± 7.32 HC: 54.6 ± 10.5	HY scale: 2.07 ± 0.39 UPDRS: PD: 31.00 ± 10.37 HC: 0.83 ± 1.27 Disease Duration: Not specified Medication: OFF	Twenty-channel EEG recorded during 5 min in resting state with eyes-closed. Two recordings were performed per patient. Cleaning was not specified.	Coherence analysis with 2 s windows with 50% overlap was carried out. Pearson's correlation was calculated to assess the relationships between coherence and disease severity. The relative and absolute PSD were calculated at 1–40 Hz. Only 14 EEG-based features were used.	DFA	A linear DFA was used. The classifier input was selected by utilizing the step-wise discriminant analysis procedure in the SPSS software package.	Cross-validation.	Accuracy = 95.24 Sensitivity = 94.74 Specificity = 95.65 PPV = 94.74 NPV = 95.65 An excessive coherence was observed in the beta and gamma bands for PD.

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[45]	Classification of PD patients with MCI vs. NC	Subjects: 36 PD with MCI and 35 PD with NC Age: PD-MCI: 61.1 ± 8.2 PD-NC: 57.0 ± 11.9	HY scale: PD-MCI: 2.1 ± 0.7 PD-NC: 1.9 ± 0.7 UPDRS: PD-MCI: 21.9 ± 8.7 PD-NC: 21.7 ± 9.4 Disease Duration: PD-MCI: 2.8 ± 2.3 PD-NC: 3.6 ± 3.6 Medication: ON	Sixteen-channel EEG recorded during 30 min at 250 Hz. The impedances were set to $Z > 100 \text{ M}\Omega$ . High-pass filter at 0.16 Hz and low-pass filter at 500 Hz were applied before pre-amplification. Wavelet decomposition and reconstruction were made. Within 60 s segments, epochs with eyes-open without any obvious artefacts were selected. Artefacts were further eliminated by ICA.	Sixty-four features were calculated from power spectrum for 4 bands. Other features from MR were calculated. The feature importance method of Mean Impact Value was used to categorize the contribution of all features.	SVM	RBF kernel was applied. The regularization parameters were identified using a “grid search”.	Five-fold cross-validation on 80% of patients. 20% for test. Additional LOO cross-validation.	Only features from EEG Train: Accuracy: 64 Sensitivity: 68 Specificity: 62 PPV: 54 NPV: 75 AUC: 66 LOO-CV: 66 In test: Accuracy: 67 Sensitivity: 67 Specificity: 67 PPV: 75 NPV: 57 AUC: 71
[46]	Classification of HC vs. patients with different types of psychological disorders.	Subjects: 25 PD and 25 controls Age: PD: 69.68 ± 8.73 HC: 69.32 ± 9.58	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Sixty-channel EEG in Oddball task at 500 Hz. A band-pass filter at 0.1–60 Hz was applied. Low variance electrodes are dropped, so 30 electrodes were selected.	Pre-processed data were segmented into 4 s epochs in case of non ERP data and original trial length is kept in case of ERP data. In total, 26 linear and non linear features including time and frequency domain features were calculated using linear SVM classifier.	SVM, LR, KNN, DT	SVM: RBF kernel. Grid search technique is used to tune the hyper-parameters	Five-fold cross validation	On PD patients vs. HC the best results for the selection of electrodes: Accuracy: 82 F1 Score: 80 Precision: 85 Recall: 82 Model not specified.
[47]	Selection of the best QEEG features to identify different levels of cognitive impairment in PD.	Subjects: 118 PD classified into 5 groups according to the severity of the disease. G1: n = 28, G2: n = 33, G3: n = 43, G4: n = 5, G5: n = 9. Age: G1 = 60.54 ± 8.75 G2 = 66.09 ± 6.65 G3 = 67.04 ± 7.94 G4 = 73.19 ± 5.29 G5 = 67.56 ± 5.51	HY scale: G1 = 1.93 ± 0.4 G2 = 2.14 ± 0.55 G3 = 2.21 ± 0.59 G4 = 2.40 ± 0.55 G5 = 2.00 ± 0.97 UPDRS: G1 = 26.00 ± 11.73 G2 = 29.55 ± 12.22 G3 = 28.74 ± 11.44 G4 = 31.00 ± 12.79 G5 = 29.00 ± 18.87 Disease Duration: G1 = 7.75 ± 5.29 G2 = 8.36 ± 7.49 G3 = 8.81 ± 5.02 G4 = 6.60 ± 3.58 G5 = 12.00 ± 6.56 Medication: ON	One-hundre-twenty-two-channel EEG recorded during 10 min in resting state. Average reference and 0.1–100 Hz bandwidth filter. Ocular artifacts were corrected and a 50 Hz filter was applied. Periods of drowsiness were removed, and the semi-automatic rejection of artifacts was performed to eliminate muscle activity. Each channel was divided into 4 s epochs. At least 20 segments were used for the analysis.	The relative and absolute spectral power were obtained for each epoch using a FFT and a 50% overlap for the delta, theta, alpha, and beta bands. A division into 5 ROI was performed. For each case, high and low electrode density were considered. A statistical dependency study with an analysis of variance and the selection of characteristics with Pearson’s correlation method was carried out.	SVM, KNN	SVM: Gaussian kernel KNN: k = 9 and the Euclidean distance as a metric.	Data were split into a training set (n = 100) and a test set (n = 18). The training set was used for 5-fold cross-validation.	SVM: Accuracy = 87 ± 3.5 KNN: Accuracy = 88 ± 2.8 Both were achieved for the relative power with low-electrode density. Groups with few patients had worse results.

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[48]	Identify patients with early PD.	Subjects: 29 drug-off early PD, 12 drug-on early PD and 22 controls Age: 62.4 ± 6.3 PD-OFF: 65.3 ± 5.4 PD-ON: 63.8 ± 5.5	HY scale: 1 UPDRS: 15.8 ± 7.5 PD-OFF: 14.3 ± 6.2 Disease Duration: Not specified Medication: ON and OFF	Nineteen-channel EEG recorded in resting state with eyes-closed at 500 Hz with additional channels for ECG, EMG, and EOG. Impedances were kept below 5 kΩ. Fast-ICA was applied to remove artifacts. Epochs with amplitude > 80 μV were rejected. More than 5 min signals were kept for each subject. A band-pass FIR filter at 0.5–45 Hz was used. Signals were segmented into 2 s non-overlapping epochs.	EEG signals were decomposed into two bands through the FIR filter. The P-Welch function was used to calculate the PSD of each channel within each epoch at 0.5–45 Hz with step size of 0.5 Hz. Models were fed with channel-frequency PSD and structured PSD. A personalized characteristic index of frequency domain was calculated for statistical analysis.	CNN, SVM, MLP	SVM: linear kernel CNN: 3 2D-convolutional layers (with 3 × 3 filters and 8, 16 and 32 neurons respectively), with layer normalization after each of them, and 2 full-connect layers (with 4576 and 40 neurons) before the ‘Softmax’ layer. Nadam optimizer with learning rate of 0.001 was used.	The dataset was obtained shuffling drug-off early PD group and HC group. It was divided into training (80%) and test (20%). The training set was used for 8-fold cross-validation.	The CNN model on the test set of structured PSD yielded: Accuracy: 99.87 ± 0.03 AUC: 99.88 ± 0.05.
[49]	Classification of HC vs. PD patients.	Subjects: 15 PD and 18 HC Age: 67.3 ± 6.5 PD: 67.6 ± 8.9	HY scale: 1.3 UPDRS: 23.3 ± 9.1 Disease Duration: 7.4 ± 4.3 Medication: OFF	Twenty-seven-channel EEG recorded during 60 s in resting-state with eyes-open at 1000 Hz. The noise was removed.	Signals were filtered into four frequency bands, by a two-way FIR filter. A general orthogonalized directed coherence was used in each band to compute directional connectivity maps, which were normalized and converted into 2D images, resized to feed the VGG-16 model. LASSO regression models were computed for 30 runs separately on latent and non-latent cases. In each run, data were randomly divided in train (25) and test (8) sets. In total, 30 results with least MSE at each run were considered as the LASSO coefficients.	VGG-16	VGG-16 with modifications: functional VGG-16, maxpooling layer (512 units), fully connected layer (512 units), fully connected layer (64 units) with ReLU activation function, and fully connected layer (2 units) with Softmax activation function. Optimizer: SGD. Learning rate: 0.01. Decay: 0.001. Batch size: 8. Loss function: binary cross entropy.	The model was tested 10 times on randomly chosen train/test partitions. The test data set was 25% of data that had been randomly selected.	After 10 random repetitions with deep transfer learning Accuracy: 99.62 Precision: 100 Recall: 99.17 F1-score: 0.996 AUC: 0.996 Latent features were correlated with five clinical indices.

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[50]	Classification of HC vs. PD patients.	Subjects: 9 PD and 9 HC Age: PD: $55.22 \pm 6.25$ HC: $52.11 \pm 4.98$	HY scale: $2.28 \pm 0.71$ UPDRS: $25.89 \pm 7.32$ Disease Duration: Not specified Medication: Not specified	Thirty-two-channels EEG recorded in resting state with eyes-open during 15 min at 1000 Hz. Impedances were kept below 5 k $\Omega$ . An online bandpass filter at 0.1–100 Hz and an offline bandpass filter at 0.1–45 Hz were used. Artifacts were removed with ICA. Segments exceeding 150 $\mu$ V were removed. ECG, PPG and RA signals were synchronously recorded. In total, 14 epochs of 60 s without artifacts were selected per subject.	Absolute and relative powers, were computed for each electrode for 4 frequency bands, as well as sensory-motor rhythm, and the ratio of alpha to theta spectrums. For feature selection, the elastic network was employed with optimal norm regularization parameters L1 and L2.	SVM	A linear kernel was used.	Nine-fold cross validation. Experiments were randomly repeated 10 times. In each one, the model parameters and feature selection were determined by inner iterations	For only EEG: Accuracy: $87.54 \pm 13.46$ Sensitivity: $86.19 \pm 15.14$ Specificity: $88.89 \pm 19.75$ The EEG of PD patients had a significant decrease in high-frequency power.
[51]	Classification of 3 diseases between them and vs. controls.	Subjects: 16 PD and mached controls Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: ON	Eight-channel EEG during 10 s in a CNV trial. Some artifacts were removed and a low-pass filter at 30 Hz was applied. Mean level and baseline corrections were performed. For each trial, 2 epochs of 512 ms were extracted.	Sixteen amplitude measures and a time measure were generated. The time measure included the post-imperative negative variation.	MLP	Input layer with 17 nodes, 1 hidden layer with 40 nodes, and 1 node for the output layer. The parameters were: gain = 1, momentum = 0.6 and learning rate = 0.9. Back-propagation was used.	LOO cross-validation.	For PD vs. controls: Sensitivity = 100 Specificity = 94 False-ve = 0 False+ve = 6 PPV = 94 NPV = 100 FAR = 6 FRR = 0
[52]	Classification of HC vs. PD patients with and without medication.	Subjects: 15 PD and 16 HC Age: PD: $63.2 \pm 8.2$ HC: $63.5 \pm 9.6$	HY scale: 2–3 UPDRS Not specified Disease Duration: $4.5 \pm 3.5$ Medication: ON and OFF	Thirty-two-channel EEG during 3 min in resting state with eyes-open at 512 Hz. The mean of the data was removed and were re-referenced to the common average. A highpass filtering at 0.5 Hz was used. The artifacts were manually examined and removed. Data were segmented into 2 s epochs.	An automated tunable Q wavelet transform was used to extract representative subbands. Five features were extracted from the subbands. The clinical significance of features are tested using the Kruskal–Wallis test.	SVM, ANN, KNN, RF, LSSVM	LSSVM: polynomial (d = 10), RBF (sigma = 0.05), Morlet (ai = 0.01), Sinc and Mexican Hat (ai = 22 and constant omega = 0.5) were tested. KNN: k = 10. RF: total number of learner was 10. SVM: Fine Gaussian kernel with automatic boxconstraint level. ANN: 10 hidden neurons.	Ten-fold cross-validation.	The results with LSSVM were: HC vs. PD OFF Accuracy: 96.13 AUC: 97 HC vs. PD ON Accuracy: 97.65 AUC: 98.56



Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[53]	Identification of the cognitive decline in PD patients.	Subjects: 20 PD patients with good cognition and 20 PD patients with poor cognition Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Twenty-one-channel EEG recorded during 15–20 min in resting state with eyes-closed. The signals were band-pass filtered at 0.16–70 Hz. For each patient, 5 visually inspected artifact-free epochs of 8.192 s were extracted.	Seven-hundred-ninety-four-time features were extracted from each channel. In addition, several clinical and spectral features were considered. The boruta algorithm was used for feature selection.	RF	The Bayesian algorithm was chosen for hyperparameter optimization. The depth of each tree and the number of trees were 1 to 100. The minimum number of samples in the leaf node and to split a node were 1 to 10 and 2 to 20, respectively.	Five runs of 10-fold cross-validation	The modeling approach (4) had the best performance. The boruta algorithm selected 1 to 3 features. None of them were clinical. Accuracy = $84.0 \pm 4.2$ F1 = $85.6 \pm 3.2$ Precision = $88.3 \pm 6.9$ Sensitivity = $83.0 \pm 2.7$ AUC = $86.8 \pm 6.0$
[54]	Classification of PD patients vs. HC.	Subjects: 25 PD and 25 HC Age: PD: $58.7 \pm 7.7$ HC: $54.6 \pm 7.7$	HY scale: 1–2 UPDRS: Not specified Disease Duration: $5.6 \pm 3.5$ Medication: Not specified	Seventeen-channel EEG recorded during 5 min in resting state with eyes-closed at 200 Hz. Impedances were kept below 5 k $\Omega$ . Signals were amplified, a notch filter at 50 Hz, and a bandpass filter at 0.5–30 Hz were applied. ICA was used. The data were re-referenced to the average reference. Signals were segmented into 5 s windows with 2.5 s overlap.	EEG segments were decomposed, using dynamic mode decomposition, into stable and unstable components at four frequency band. By Pearson correlation, stable brain network, unstable brain network, and inter-connected brain network were constructed and thresholded separately. Traditional brain network was also constructed. Topological attributes were extracted.	SVM, BN, RF, SGD, KNN, Adaboost, RT, bagging, SL, vote methods	Not specified	Ten-fold cross-validation	The average results for all classifiers: Using the stable brain network attributes Precision: 89.6 Recall: 89.5 AUC: 90.8 Using traditional brain network attributes Precision: 85.4 Recall: 85.0 AUC: 85.4
[55]	Detection of Turning Freezing in PD patients.	Subjects: 6 PD Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: OFF	Fifteen-channel EEG recorded during Timed Up and Go tasks. The EEG was segmented to 1 s epochs associated with normal turning or turning freezing, which resulted in 204 s of each of them. A band-pass filter at 0.5–40 Hz was used. Artifacts were removed. Z-transformation was applied to normalize the signals.	Two parameters were extracted from the theta, alpha, low beta, and high beta spectral bands using S-transform. These were the maximum amplitude for each band and the sum of amplitude of each band. ICA-EBM was considered for source separation.	BNN	Three-layer (input, hidden, and output layers) feed-forward structure. Either 4 (for 15 channels) or 7 (for 4 channels) hidden nodes were considered.	Data were randomly split into 50% for training and 50% for test sets.	Best results with ICA, 4-channels and 7 hidden nodes. Train: Accuracy = 86.8 Sensitivity = 85.8 Specificity = 88.0 Test: Accuracy = 86.2 Sensitivity = 84.2 Specificity = 88.0 AUC = 0.9296

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[56]	Detecting the occurrence of GIF	Subjects: 4 PD Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: OFF	Thirty-two-channel EEG during a structured series of Timed Up and Go tasks at 512 Hz. Only data from 9 electrodes positioned in locations of interest were processed. Average of 2 ear lobes electrodes was taken as reference. Data were segmented into 1 s windows and filtered using a nonlinear IIR band-pass filter with a cut-off frequency lower than 1 Hz and higher than 50 Hz to remove artifacts. In total, 122 EEG samples were collected, associated to Good Start (61) and Gait Initiation Failure (61).	Welch's method with a 256 points FFT with 55% overlapping was used to analyze four frequency sub-bands. Power spectra density and centroid frequency of each band were calculated.	BP-NN	Two-layer feed-forward neural network with 10 hidden nodes. Activation function of the hidden layer was tan sig. Levenberg Marquardt's algorithm with early stopping was used.	In total, 50 times runs. Data was randomly divided into training 34%, validation 33%, and testing 33%.	The best performance of the classification system was achieved with a combination of nine channels: Sensitivity: 84.27 Specificity: 85.02 Accuracy: 84.80
[57]	Classification of PD patients with medication, without medication, and controls.	Subjects: 10 PD and 12 controls Age: PD: (40–80) HC: (40–80)	HY scale: 1–2 UPDRS: Not specified Disease Duration: Not specified Medication: ON and OFF	Sixty-one-channel EEG recorded during 2 min in resting state followed by 140 s of intermittent photic stimulation with eyes-closed. In total, 10 electrodes were selected. Band-pass filter at 0.5–50 Hz and a 60 Hz notch filter were applied. Average reference was used. The artifacts were removed. The impedance remained below 5 kΩ.	The last 10 s stretch of stimulation was divided into 20 segments of 0.5 s which were used to calculate the partial directed coherence for 6 bands (delta, theta, alpha, beta, gamma1, and gamma2). A total of 60 features were calculated. In total, 19 features were selected using GA with a population of 20 individuals, 20 generations, a crossover probability of 0.6, and a mutation probability of 0.03. Both approaches were considered as inputs.	BN, NB, MLP, SVM, J48, RT, RF, ELM, mELM	MLP: hidden layers 1 and 2, learning rate 0.3, momentum 0.2, iterations 500. SVM: polynomial kernel (exponent = 1 to 5) RBF kernel (gamma = 0.25 and 0.5). RF: Trees 10 and 50. ELM: 100 neurons in the hidden layer and sigmoid kernel. mELM: 100 neurons in the hidden layer, dilatation, and erosion kernels.	K-fold cross-validation.	The best model was RF with 50 trees. All features: Accuracy = 99.22 19 features: Accuracy = 98.09

**Table 3.** *Cont.*

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[58]	Classification of PD patients vs. controls using 6 emotional stimuli.	Subjects: 20 PD and 20 controls Age: PD: (40–65) HC: (40–65)	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Fourteen-channel EEG recorded during emotional stimuli. The signals were segmented into 10 s epochs with overlapping of 75%. Then, pass-band elliptic filters were used to obtain the alpha, beta, and gamma bands.	For each band, 6 features were calculated: Entropy (EN), Energy-Entropy (EEN), Teager Energy-Entropy (TEEN), Spectral Entropy (SEN), Spectral Energy-Entropy (SEEN), and Spectral Teager Energy-Entropy (STEEN).	PNN, KNN, SVM	KNN: k in the range 1 to 10. PNN: used exponential activation function with $\sigma$ ranged from 0.55 to 0.65.	Not specified.	The best accuracy for each emotion: Happiness: 99.59 with SVM and TEEN Sadness: 90.81 with KNN and EN Fear: 95.07 with KNN and SEEN Anger: 91.42 with SVM and SEEN Surprise: 94.53 with KNN and SEEN Disgust: 88.18 with SVM and SEEN.
[59]	Classification of PD patients vs. controls using 6 emotional stimuli.	Subjects: 20 PD and 20 controls Age: PD: (40–65) HC: (40–65)	HY scale: Not specified UPDRS: Not specified Disease duration: Not specified Medication: Not specified	Fourteen-channel EEG recorded during emotional stimuli. The signals were segmented into 10 s epochs with overlapping of 75%. Then, pass-band elliptic filters were used to obtain the alpha, beta, and gamma bands.	For each band, 4 features were calculated: Entropy (EN), Energy-Entropy (EEN), Spectral Entropy (SEN), and Spectral Energy-Entropy (SEEN).	PNN, KNN	KNN: k in the range 1 to 10. PNN: exponential activation function with $\sigma$ ranged from 0.55 to 0.65.	Not specified.	The best accuracy for each emotion: Happiness: 96.8 with PNN and SEEN Sadness: 90.81 with KNN and EN Fear: 95.07 with KNN and SEEN Anger: 88.65 with KNN and EN Surprise: 94.53 with KNN and SEEN Disgust: 87.43 with KNN and EN.
[60]	Classification of 3 diseases and controls through 4 methodologies.	Subjects: 15 PD and 16 controls Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Twelve-channel EEG recorded during approximately 250 ms. Cleaning and task were not specified.	Features selected through 4 methodologies: 1. The maximum amplitude and frequency of the FFT in alpha, beta, theta, and gamma bands. 2. Amplitudes of the first 8 peaks of the FFT and their frequency values up to 30 Hz. 3. The FFT of the EEG signal below 30 Hz. 4. The maximum amplitude and frequency of the FFT in each band were applied to the combination of the 12 electrodes.	ANN	Feed-forward architecture with 16 nodes in the hidden layer. Activation functions were logsig and tansig. The training algorithm was Levenberg-Marquardt.	In total, 75% training, 12.5% validation and 12.5% test. The sets were balanced in terms of pathologies that contained each one.	Bad results were expressed in terms of the correlation coefficient between target and predicted values.

Table 3. Cont.

Ref.	Objective	Evaluation of the Quality of the Dataset			Data Pre-Processing	Models	Evaluation of the Models Used	Validation	Results
	Objective	Participants	Stage PD	EEG Pre-Processing	Features		Model Parameters		Best Results
[61]	Classification of RBD patients vs. controls. Some patients were eventually diagnosed with PD and dementia.	Subjects: 118 RBD and 74 controls. 14 RBD became PD. No direct patient data. Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Fourteen-channel EEG recorded in resting state with open-eyes periods followed by closed-eyes periods. Only eyes-closed sequences were considered. The EEG was recorded when the patients were RBD. A band-pass filter at 0.3–100 Hz and a notch filter at 60 Hz were applied. Artifacts were removed.	Using a sliding window of 1 s, 148 spectrograms per subject were generated (of 20 s of duration each) using only the FFT amplitude bins in the band 4–44 Hz. The spectrograms were centered and normalized to unit variance for each frequency and channel.	CNN, RNN	CNN: 4 hidden-layer convolutional. Dropout, max-pooling layers, and cross-entropy loss function were used. RNN: with LSTM and GRU, with 3 cells which 32 units each. Dropout was used.	LOO cross-validation. For training, the dataset was balanced by random replication preserving the distribution of the subjects.	The results for controls vs. PD: CNN: Accuracy = $79 \pm 1$ AUC = $0.87 \pm 0.1$ RNN: Accuracy = $81 \pm 1$ AUC = $0.87 \pm 0.1$ In RNN, there was no difference between LSTM and GRU.
[62]	Classification of PD patients with RBD vs. controls.	Subjects: 14 RBD with PD and 14 controls. Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Fourteen-channel EEG recorded in resting state with open-eyes periods followed by closed-eyes periods. Only eyes-closed sequences were considered. The EEG was recorded when the patients were RBD. A band-pass filter at 0.3–100 Hz and a notch filter at 60 Hz were applied. Artifacts were removed.	Several spectrograms were computed to extract temporal series of power for each electrode and band (10 bands in total). The use of 4 s and 1 s spectrogram windowing was explored.	RNN	ESN layer with 3000 nodes, least-squares regularization. Spectral radius ranged from 0.5 to 2.	For each parameter set, 50 runs were carried out with random and balanced training (90%) and test (10%) sets.	The best performance was obtained with 1 s. Test set: Average accuracy = 85
[63]	Classification of PD patients vs. controls.	Subjects: 30 PD and 30 controls. Age: PD: (50–70) HC: (50–70)	HY scale: 1–1.5 UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Two-channel EEG recorded for 30 min for the flexion and extension of the wrist. Cleaning was not specified.	EEG: Shannon entropy, Lyapunov, and inverse Lyapunov exponent were calculated. EMC: power, standard deviation, root mean square, variance, waveform length, modified median, and mean frequency.	MLP	Back Propagation was used as the learning algorithm and “trainlm” was used as the training function. Sigmoid transfer function was used for the hidden layer.	The dataset was divided into training 70%, validation 15%, and testing 15%.	MLP with inputs: EEG: accuracy = 62 EMG: accuracy = 73 EEG + EMG: accuracy = 98.8
[64]	Classification PD and HC	Subjects: 16 PD and 15 controls. Age: PD: $62.6 \pm 8.3$ HC: $63.5 \pm 9.6$	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: OFF	Forty-channel EEG during 2 min in resting state at 512 Hz. Signals were segmented into patches of 512 time samples.	EEG patches of 512 samples were used as input. Channels were considered independently.	ANN	Twelve layers with 512, 512, 128, 128, 64, 64, 32, 32, 16, 16, 8, and 2 (output) units. The conjugate gradient backpropagation with Polak-Ribière updates was used for updating the weights, and biases of the network. Grid search was used for hyper-parameters selection.	In total, 20% for testing and the rest was used for training and validation.	The Oz/P8/FC2 channels were selected and used with majority voting. Results on test set were: Accuracy: 98 Sensitivity: 97 Specificity: 100

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[65]	Classification of PD patients with ON medication vs. OFF medication.	Subjects: 28 PD Age: 69.75 ± 8.43	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: ON and OFF	Sixty-four-channel EEG recorded during 12 min in the oddball task. In total, 60 electrodes were selected during the period without stimuli. A band-pass filter at 13–30 Hz was applied. Non-brain activities were subtracted.	The average spectral power over all the frequency ranges was calculated and added together. The differences between OFF and ON medication data were compared. In total, 12 channels were selected. A preprocess with delay ( $\tau = 12$ ) and embedding dimension ( $m = 3$ ) were applied. Three sets of 12 time-series (non-delayed, $\tau$ -delayed and $2\tau$ -delayed) were generated.	EEGNet, HNet, DGHNet	ADAM optimizer with learning rate = $10^{-4}$ for 500 epochs was used. HNet: 3 layers (Convolutional, LSTM, LSTM) with 28 filters, 14 and 2 neurons, with sigmoid, sigmoid, and Softmax activation. DGHNet: 2Dconvolutional as input layer with 4 filters and exponential linear activation. Values of $m = 4$ and $\tau = 3, 6, 9, 12,$ and $17$ were trained. 1Dconvolutional with exponential linear activation. LSTM with 2 outputs and softmax activation.	Intra-patient: the data was randomly shuffled and split into 90% training and 10% test. The training dataset was divided into 85% training and 15% validation. Inter-patient: 2 patients with ON and OFF medication were randomly left out for the test. The remaining patients were partitioned at 85% training and 15% validation.	The results of HNet and DGHNet were similar (accuracy: 99.74 vs. 99.22). DGHNet was considered better for containing fewer parameters. Test: Accuracy = 99.22 Sensitivity = 98.98 Specificity = 99.46 MCC = 98.44 F1 = 99.24
[66]	Classification of PD patients vs. HC	Subjects: 9 PD and 9 HC Age: Not specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Twenty-channel EEG recorded in 3 epochs of 1 min during externally paced bilateral, cyclical foot in the sitting position at 1200 Hz.	Kolmogorov Complexity, Sample Entropy, and Approximate Entropy were extracted for each channel. Different subsets of electrodes were considered, and their associated features were used as input.	MLP	Five hidden layers chosen by trial and error method.	Ten-fold cross validation	The best results were obtained for the channels Fz, F1, and F2: Accuracy: 97.5 Precision: 100 Sensitivity: 96.7 Specificity: 100 AUC: 0.978
[67]	Selection of the best classifier of PD vs. controls using the minimum number of HOS features.	Subjects: 20 PD and 20 controls. All right-handed. Age: PD: 59.05 ± 5.64 HC: 58.10 ± 2.95	HY scale: 1: n = 2; 2: n = 11; 3: n = 7 UPDRS: Not specified Disease Duration: 5.75 ± 3.52 Medication: ON	Fourteen-channel EEG recorded during 5 min in resting state with eyes-closed. Threshold technique at 80 $\mu$ V. A band-pass filter at 1–49 Hz was applied. 2 s epochs with 50% overlap were considered.	For each epoch, a total of 13 HOS characteristics were calculated. The Student's t-test was also obtained to determine the importance of the characteristics.	DT, KNN, FKNN, NB, PNN, SVM	FKNN: Euclidean distance, $m = 1.24$ and $k = 3$ . KNN: $k = 2$ and Euclidean distance. PNN: exponential activation and $\sigma = 0.284$ . SVM: polynomial (orders 2 and 3), RBF, and linear kernels.	Ten-fold cross-validation. The characteristics were added one by one to each classifier until maximum precision was achieved.	The best model was SVM with RBF kernel: Accuracy = 99.62 ± 0.58 Sensitivity = 100 ± 0.0 Specificity = 99.25 ± 0.53 Precision = 99.38 ± 0.47 F1 Score = 0.98 ± 0.05

Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used			Results
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[68]	Identification of the onset of freezing of PD patients during walking	Subjects: 26 PD Age: 69.8 ± 8.41	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Four-channel EEG recorded in periods of 1–2 h per patient at 500 Hz during a structured series of video-recorded Timed Up and Go tasks. Only data from 10 patients and differential channels O1-T4 and P4-T3 were used. Epochs of 1 s from individual freezing events were taken. Band-pass at 0.5–60 Hz, and band-stop at 50 Hz butterworth IIR filters were applied. An additional threshold filter was applied.	Data were divided into normal, onset, and freezing with 40 samples per subject and group. Discrete wavelet transform was used to calculate wavelet entropy for 5 frequency bands, and total wavelet entropy. The model was fed with the wavelet entropy of 3 bands in total for channel O1-T4, and all bands and total for channel P4-T3, separately and in combination.	BP-NN	Three layers with 4–7 hidden nodes depending on the number of inputs dimension and the number of training pairs. Levenberg Marquardt algorithm was used. Activation function was Tangent Sigmoid.	Twenty runs for each feature. Data was divided in training 56%, validation 25%, and test 19%.	The results for Normal vs. Onset for P4-T3 were: Accuracy: 76.6 ± 3.4 Sensitivity: 74.2 ± 6.8 Specificity: 78.9 ± 7.3 Normal vs. Freezing for O1-T4 and P4-T3: Accuracy: 73.9 ± 2.8 Sensitivity: 71.2 ± 6.1 Specificity: 77.2 ± 4.7
[69]	Predicting transition to FOG from normal walking	Subjects: 26 PD Age: 69.8 ± 8.41	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	Four-channel EEG recorded in periods of 1–2 h per patient at 500 Hz during a structured series of video-recorded Timed Up and Go tasks. Data from 10 patients without significant artifacts were selected and differential channels O1-T4 and P4-T3 were used. Epochs of 1 s were taken. Band-pass at 0.5–60 Hz, and band-stop at 50 Hz butterworth IIR filters were applied. Ocular and muscular artifacts were removed.	Data were divided into normal, transition, and freezing with 40 samples per subject and group. Discrete wavelet transform was used to calculate total, global and centroid frequency wavelet cross spectrum for 5 bands, wavelet cross frequency energy ratios por specific pairs of bands. Other statistical features were computed. In total, 131 features were selected to feed the model separately and in combination.	MLP, KNN	MLP: 3 layers with 8 to 12 hidden layer neurons. The Levenberg Marquardt algorithm was used. Error goal of 0.01. KNN: 15 to 40 nearest neighbors based on the Euclidian distance.	Twenty runs for each feature and their combination. Data was divided into training 56%, validation 25%, and test 19%.	Using MLP with statistical features: Sensitivity: 75.47 Specificity: 71.47 Accuracy: 73.47 For kNN with all features were: Sensitivity: 87.25 Specificity: 70.00 Accuracy: 52.75



Table 3. Cont.

Objective		Evaluation of the Quality of the Dataset			Data Pre-Processing	Evaluation of the Models Used		Results	
Ref.	Objective	Participants	Stage PD	EEG Pre-Processing	Features	Models	Model Parameters	Validation	Best Results
[70]	Classification of HC vs. PD patients through different emotional states.	Subjects: 5 PD and 5 controls Age: PD: 45–65 HC: 20–25 Age for each subject were specified	HY scale: Not specified UPDRS: Not specified Disease Duration: Not specified Medication: Not specified	One-channel EEG recorded while watching happy, sad, and neutral videos, and during meditation in resting state with eyes-closed. The EEG collecting device produced power values for eight frequency bands, and attention values.	Benjamini-Hochberg corrected F-tests and t-tests were applied on attention level, average power of frequency bands, and on absolute value of first and second order derivatives. An initial model was fed with first and second order derivatives for each pair emotion-band. Meditation was not used. The three best pairs were combined in the final model. Input features were preprocessed to have zero mean and unit variance.	MLP	Initial model: 1 hidden layer with 4 neurons. Final model: 2 hidden layers with 4 and 2 neurons, respectively. Adam optimizer, Log-loss loss function, and activation function tanh were used in both.	Five-fold cross validation	The results of the final model were: Accuracy: 0.965 F1 score: 0.976 Recall: 0.970 Precision: 0.955
[71]	Classification of patients with neurological diseases vs. controls.	Subjects: 31 PD and 264 controls. Age: PD: 56.62 ± 12.32 HC: 49.51 ± 12.54	HY scale: 1–3 UPDRS: 43.44 ± 15.53 Disease Duration: Not specified Medication: Not specified	Nineteen-channel EEG recorded during 5 min in resting state with eyes-closed. The impedances were kept below 5 kΩ. A high-pass filter at 0.15 Hz and a low-pass filter at 200 Hz were used. A band-pass filter at 2–44 Hz was applied. Artifacts were manually removed. Channels were divided into ROI.	The power spectrum was calculated for each subject and 5 frequency bands (delta, theta, alpha, beta, and gamma) were considered for each ROI.	SVM	The default settings were used as the running parameters.	Ten-fold cross-validation. The distribution of the patients was kept. Controls with obesity were used to validate the model.	The results for controls vs. PD: Accuracy = 94.34 ± 1.81 Sensitivity = 0.93 ± 0.02 FPR = 0.11 ± 0.01 ROC = 0.95 ± 0.02 MAE = 0.07 ± 0.02 RMSE = 0.16 ± 0.02

Acronyms: QEEG—Quantitative Electroencephalogram, EEG—Electroencephalogram, PD—Parkinson’s disease, HY—Hoehn-Yahr scale, UPDRS—Unified Parkinson’s Disease Rating Scale, EEG—electroencephalogram, FOG—freezing of gait, DTF—directed transfer function, H-COG—high cognition, L-COG—low cognition, inter-COG—intermediated cognition, AOB—auditory Oddball, BNA—Brain Network Analytics, SPWVD—smoothed pseudo-Wigner Ville distribution, TFR—time-frequency representation, NC—normal cognition, ERP—event-related potential, ECG—electrocardiogram, PPG—photoplethysmography, RA—respiratory, LSSVM—least square support vector machine, EW—Effective walking, SGD—steepest gradient descent, SL—simple logistic, GIF—Gait Initiation Failure, BP-NN—Back Propagation Neural Networks, FFT—Fast Fourier Transform, ROI—regions of interest, SVM—Support Vector Machine with C regularization constant and  $\sigma$  width of the kernel, KNN K—Nearest Neighbors with k being the number of nearest neighbors considered, HC—healthy controls, LOO—leave-one-out, RBD—REM behavior disorder, REM—Rapid eye movement, EOG—electrooculogram, RF—Random Forest, MCI—mild cognitive impairment, PLI—Phase lag index, DT—decision tree, LR—logistic regression, LASSO—Least Absolute Shrinkage and Selection Operator, RBF—radial basis function, AUC—Area Under the Curve, MLP—Multilayer Perceptron, PPV—positive predictive value, NPV—negative predictive value, FAR—false alarm rate, FRR—False reassurance rate, CNV—contingent negative variation, SFAM—Simplified Fuzzy ARTMAP, PSFAM—Probabilistic SFAM, IPSFAM—integrated PSFAM, PNN—Probabilistic neural network with  $\sigma$  being the smoothing parameter, CNN—Convolutional Neural Network, RNN—Recurrent Neural Network, LSTM—Long-Short Term Memory Network, ICA-ENM—independent component analysis by entropy bound minimization, BNN—Bayesian neural networks, GA—Genetic algorithms, BN—Bayes net, NB—Naïve Bayes, RT—random tree, ELM—extreme learning machine, mELM—morphological ELM, GRU—Gated-Recurrent Unit, ESN—Echo State Network, EMG—Electromyogram, RMSE—Root-mean-square error, HNet—Hybrid Network, DGHNet—Dynamical system Generated HNet, MCC—Matthews Correlation Coefficient, ANN—Artificial Neural Network, PCA—principal component analysis, FC—agglomerative feature clustering, CSP—Common Spatial Patterns, FPR—false positive rate, ROC—Receiver Operating Characteristic, MAE—mean average error, PSD—power spectrum density, DFA—Discriminant Function Analysis, HOS—higher order spectrum, CFS—correlation-based feature selector, FKNN—fuzzy KNN with m being the fuzzy strength parameter. Here n stands for the number of patients.

1. Evaluation of the quality of the dataset. It was evaluated if the articles considered balanced datasets, and if the samples were statistically equivalent, by means of the number of subjects in each group and the demographic data of the patients. Other clinical parameters related to PD progression were also evaluated, such as the Hoehn-Yahr (HY) scale, UPDRS, years of disease, and whether the data collection was recorded with active dopaminergic medication. These parameters provided information about the quality of the dataset, allowing for assessing the performance results of the model. Finally, parameters related to the recording of EEG signals were evaluated, such as the number of channels, duration of the test, and type of test performed, which provided information about the quality of the signal recording process. Regarding Table 3, this point corresponds to the Participants column, Stage PD column, and part of the EEG Pre-processing column.
2. Data pre-processing. The cleaning protocol of the EEGs and the extraction of features were analyzed. The cleaning protocol of the EEG is a process that is sometimes omitted. Moreover, there is no gold standard defined, and the great variety of techniques usually considered produce different modifications in the EEG signals. Thus, the cleaning of the EEG was evaluated in each article to verify the impact of this pre-processing on the results of the models. With regard to the extraction of features from the EEG signals, it should be taken into account that the features introduced as inputs in the models play a transcendental role in the ML techniques, so they were collected to extract those used most frequently and those for which better results were obtained. The information associated with this point was specified in part of the EEG Pre-processing column, and the Features Column, of Table 3.
3. Evaluation of the models used. The type of model used together with its architecture and its training and validation methods were examined. This set of parameters was considered to assess which models obtained the best results depending on the objective of the article. More specifically, the analysis of the validation process carried out allowed for evaluating the quality of the results and provided a more objective assessment of the scope of the predictive results of the model. Within Table 3, this information corresponds to the Models, Model Parameters, and Validation columns.

The aspects exposed above combine both clinical and computational points of view, providing, together, an analysis of the three fundamental steps that should be considered in a ML problem. In addition, to facilitate subsequent analysis, Table 3 included a row in the header, specifying the correspondence between these three and the columns, an additional column indicating the objective of each study, and another column with the most relevant results of each model and the metrics used.

Next, a global analysis for each of the three key points introduced in Table 3 was provided: dataset quality, data pre-processing, and model evaluation. Regarding the objective of the selected articles, it could be noticed that 30 of the studies included in this review covered the problem of classification between patients with PD and controls, whereas the remaining articles considered diverse topics: six detected alterations in gait, four classified cognitive impairments, two selected the features to classify cognitive impairment and patients with PD, and one distinguished between patients with medication vs. patients without medication. The remaining 16 articles were not included in Table 3, despite passing the exclusion criteria, since their objective was not completely related to the study of PD with EEG. Within this group of articles, five classified emotions [72–76], four identified sleep disturbances [77–80], five used DBS or neurostimulation [81–85], one combined EEG and EMG features [86], and one classified mental tasks [87]. In conclusion, it can be appreciated that the diagnostic problem of PD is the main objective of this type of study. The subsequent evaluation was carried out in the 43 articles included in Table 3.

### 3.2.1. Assessment of the Quality of the Dataset

The evaluation of the quality of the dataset was performed through the analysis of the number of subjects that contained each class, the clinical parameters associated with the progression of the disease, and the EEG recording parameters.

It was observed that 69.77% of the selected studies (30 articles) used balanced classes, that is, the classes contained a similar number of subjects. The criterion that was chosen to evaluate when the classes were unbalanced was that the difference between each pair of classes was greater than 25% of the class with the largest number of samples. According to this criterion, 16.28% of the studies did not consider balanced classes, and the articles [35,40,55,56,68,69] only had a single class. A training set with unbalanced classes can lead to prediction errors and poor data generalization. It should be noted that for this selection, the total number of subjects per class was taken into account in those studies that used a mixture of patients with different diseases. The mean number of subjects for the balanced sets was approximately  $23.10 \pm 16.74$  in each class (this result was calculated for the PD patients). It can be appreciated that among the studies with a balanced dataset, only [41] exceeds 50 subjects in one of the classes, whereas the mode was 20 subjects per class (considered in nine articles). In addition, to verify that the classes could be statistically compared, it was convenient to verify that the subjects exhibited the same demographic characteristics. Of the 30 articles that considered balanced classes, only 18 of them specified the mean age of the patients included in the study. The mean of these groups was  $63.37 \pm 4.22$  years, which is approximately the age at which this disease usually begins. Of the remaining articles belonging to this group, eight indicated the age range of the subjects, contained in (40–80) years. The articles [43,51,53,60,62,66] did not specify the age of the subjects included in the study.

The state of the disease is relevant information to evaluate the quality of the dataset because it can influence the performance of the model and therefore affect the classification problem, since patients with PD in more advanced stages of the disease who were monitored without their habitual take of dopaminergic medication may be easier to distinguish from the controls than a group of PD patients in the early stages of the disease who took their dopaminergic medication. Nevertheless, not all articles contained the state of the disease of the patients. From Table 3 it was observed that 32.56% of the selected articles did not provide any information about the disease status or the medication, 13.95% of the studies only indicated the medication, 6.98% only specified the affection through the HY scale, and 11.63% of the articles provided little information. Therefore, only 34.88% of the selected articles included all the information related to the patients' condition and their medication. For the dopaminergic medication, it was found that 51.16% of the articles did not report the status of the medication during the EEG recording, 13.95% recorded the EEG in the ON state, 16.28% recorded the EEG in the OFF state, and 18.60% recorded the EEG in ON and OFF states.

In relation to the EEG recording, the parameters specified in Table 3 were: the number of EEG channels, the EEG recording length, and the test performed. A high density of electrodes has greater spatial resolution and therefore provides more information about the global state of the brain by increasing the data contained in the EEG. The number of electrodes was specified in all articles with a mean value of  $43.34 \pm 62.18$  electrodes. As can be appreciated, there was great heterogeneity in the density of electrodes, with a mode of 14 electrodes considered in eight studies. There were studies (such as [35,41,63,68–70]) that used EEG with less than five channels, but it should be noted that three of them detected gait disturbances, two diagnosed PD in combination with EMG, and [70] used one channel with emotional stimuli. Regarding the duration of the EEG recording, the mode was 5 min. Moreover, the tendency was to divide the signals in windows of uniform length, so that larger recordings provided larger datasets and allowed some studies to eliminate those EEG segments that had defects or could disturb the analysis. Finally, different types of EEG tests, such as resting state tests, stimulation of emotions, or motor activation tests, provide EEGs with different properties and therefore are not comparable. In the articles selected

for this review, the resting state tests predominated, as they were considered in 22 articles, followed by Timed Up and Go tasks (present in 5 articles), and tests that require a physical response to a stimulus (Oddball and Visual Go/No-Go in 4 studies).

### 3.2.2. Data Pre-Processing

Within data pre-processing, two key points may be distinguished: the EEG cleaning protocol and the feature extraction. It could be seen from the information in Table 3 that the cleaning process was very heterogeneous among the selected articles, which may be due to the lack of a standard EEG cleaning protocol. Actually, 27.91% of the studies did not specify the EEG cleaning process, 23.26% carried out little pre-processing in the EEG data, and 48.84% of the articles used artifact-free EEG signals. This makes it difficult to evaluate the dataset and assess how this pre-processing affects the EEG signals since alterations in the signals can modify essential aspects and lead to a false diagnosis. A more precise evaluation of these aspects is shown in the discussion section.

The articles also showed great heterogeneity in the features extracted from the EEG signals. Spectral features predominated (they were considered in 88.37% of the articles), and only [29,31,42,64] used signal segments as input data to the model. Although the studies used a wide variety of spectral features, the most common procedure consisted of decomposition into frequency bands.

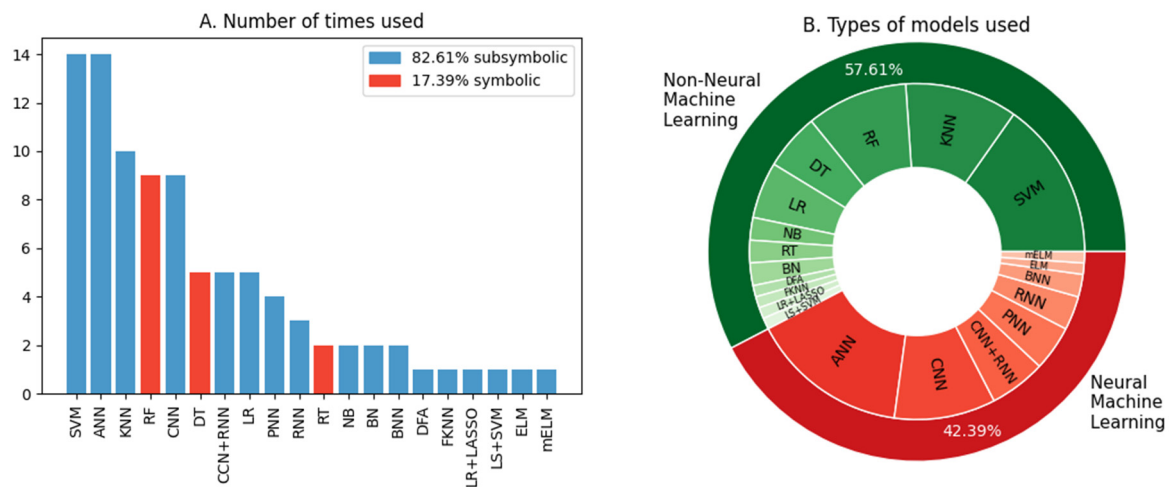
### 3.2.3. Evaluation of the Models Used

As shown in Table 3, one of the most notable characteristics of the selected articles had to do with the variety of models used. A brief description of these models can be found in Table 1 of [88]. The number of models used exceeded the number of articles selected. This was due to the fact that 18 articles, corresponding to 41.86%, made comparative studies between various models, whereas the remaining articles used a single model. Within the latter group, MLP was the most used model, being considered in five articles, followed by CNN used in four studies, and SVM and RF, utilized in three articles each.

For the complete set of selected studies, Figure 3A shows a bar chart with the models used and the number of times they appeared in the articles, differentiating those models that used symbolic processing (in red) and subsymbolic processing (in blue). It was taken into account that J48 is a model based on DT, VGG-16 is a model based on CNNs, and that SGD, AdaBoost, bagging, and vote method are training and optimization methods and therefore do not belong to the group of models used. Moreover, SL was incorporated into the group of LR. The ANN group contains both the unspecified ANN (since the corresponding articles suggested they were MLP networks) as well as BP-NN and FF-NN. As expected, since the dataset was made of time series, those models with subsymbolic processing predominated, being considered 82.61% of the time, whereas RF, DT, and RT were the only models utilized within the symbolic processing group. To emphasize these differences, it should be pointed out that whereas 17.39% of the articles considered models with symbolic processing, only two articles used them exclusively. Taking into account the previous groups, Figure 3A shows that the most used models in the articles included in this review were SVM and ANN, considered by 14 studies, followed by KNN, used by 10 studies, and RF and CNN models, which appeared in nine studies. The acronyms utilized were defined in the description in Table 3.

The importance of Neural Machine Learning, which constituted 42.39% of the models used, although the most widely used models were SVM, a non-neural ML technique, and ANN in the same proportion, stand out. Deep Learning (DL) techniques are of special importance. Deep Neural networks are networks with more complex architectures that allow greater abstraction levels at the expense of increased computational power. So, the development of these models was expected to be concentrated in recent years. In fact, since 2019, we found 28 articles, of which we discarded [46,54] for not providing enough information about the model used. Of the remaining 26 articles, 57.69% used

Neural Machine Learning models, predominating the use of CNNs, present in 34.62% of the studies.



**Figure 3.** (A) models and the number of times they are considered in the articles, indicating their belonging to either the symbolic or subsymbolic processing group. (B) Pie chart of the types used, specifying the non-neural ML techniques and those belonging to the ANN subgroup.

Regarding the architecture of the models, it should be noted that, although it was specified for most of them, the degree of specificity of this information showed great variability among the articles. Actually, for the most complex models, parameters such as the number of units in each layer, the activation function used, the optimizer used, the learning rate value, or the loss function chosen, were specified in a few of the articles, and the associated information was not always uniform for all the models within every article. Moreover, the absence of a baseline complicated the comparison between the different studies. Both points made it difficult to provide a precise assessment of the architecture, and so of the Model Parameters column. However, it is worth emphasizing that it stood out the use of the Adam optimizer, the ReLU activation function, and Softmax for the last layer. Within the assessment of the training and validation phase, it could be observed that 24 of the selected articles (55.81%) used the K-fold cross-validation technique (two of which were LOO), and 16 of the selected articles considered a division in separate sets for training/validation/testing according to percentages fixed in each study. In particular, for this last set of articles, the mode was to use around 70% of the data to train the model, considered in six of them. The remaining articles [37,58,59] did not specify the methodology used.

To evaluate the results obtained by the selected articles, the accuracy metrics were considered and used in 90.70% of the articles, followed by sensitivity, used in 69.77% of the studies, and specificity, used in 46.51% of the articles considered in Table 3. For the 31 articles that covered the diagnostic problem of PD, it was found that the models with at least two metrics greater than 90% were: refs. [51,64,66,70] using ANN with values over 97% for sensitivity and precision, refs. [38,39,49] considering CNN with values over 99% for accuracy, precision and sensitivity ([38,39] were based on the same study), refs. [29,30] using CNN + RNN with values over 93% in accuracy, sensitivity, and precision, refs. [67,71] utilizing SVM with accuracy and sensitivity values over 93%, and [44] using DFA with accuracy, specificity and sensitivity values over 94%. For these articles, the validation method carried out was evaluated to obtain information about the ability of the model to generalize the result in blind tests. It was found that the predominant method was cross-validation for 9 of the 12 articles, whereas [29,49,64] split the data into training, validation, and test. The models with the best results in the problem of classifying patients with PD and controls were CNN, and the group of ANN models, found in articles [38,39,51,59,64,66,70].



## 4. Discussion

Parkinson's is a disease mainly characterized by motor dysfunctions that affect the quality of life of patients. The development and application of ML techniques to the analysis of EEG associated with PD is a major initiative in PD research as it represents an affordable and accessible technique that may help to make an early diagnosis. As can be seen in the PRISMA diagram in Figure 1, the selection process carried out on this topic resulted in a total of 59 articles. The year and country of publication of the selected articles are displayed in Figure 2. According to that information, it can be noted that the interest in this field has shown to be global (predominating the development in Asia). Furthermore, such interest has increased in recent years, probably caused by the greater amount of available data and by the growth of computational power, which allows for the use of more complex and advanced models.

For a deeper analysis of the content of each article, three key points were evaluated within ML techniques: 1. the quality of the dataset, by means of the clinical parameters and the recording parameters of the EEG signals; 2. data pre-processing, through the cleaning protocol and the extraction of features; 3. the evaluation of the models used specifically the type of model, its architecture, and the training and validation methods. These points are summarized in Table 3.

### 4.1. Quality of the Dataset

Regarding the quality of the dataset, 69.77% of the studies worked with balanced datasets with an average number of  $23.10 \pm 16.74$  subjects for each class. The use of balanced classes in training is important to ensure reliable results since unbalanced training classes tend to favor the majority class and can lead to skewed accuracy metrics. Among the studies that used balanced classes, those with 20 subjects per class predominated. This may provide a measure of the adequate number of subjects, an especially important piece of information in the case of EEG data, as they depend on the availability of the patients, and therefore they may be quite complicated to obtain.

Furthermore, the EEGs of the subjects change with age, and hence, it should be corroborated that the classes exhibit similar demographic data so that the sets are statistically comparable. This also applies to the state of the disease of PD patients, since patients in more advanced stages of the disease are easier to distinguish from healthy controls than patients in the early stages of PD. It is striking that only 34.88% of the articles specified such information. This lack of specificity may be a limitation in many studies since the differences between the state of the disease of PD patients and their medication status may influence the results of the classification problem. This was reflected in [65], which classified PD patients in the ON state vs. OFF state of medication, and in those studies that evaluated cognitive impairment [32,34,45,47,53]. Consequently, those articles that did not specify the clinical setting of the patients [37,40,42,43,46,51,53,55,56,58–62,64–66,68–70], nor the status of the medication in the patients [32,33,35,41,50,54,63,71] were excluded from the evaluation process of the current state of diagnosis of PD, because they did not allow the objective evaluation of the results of the models used. These exclusions, based on quality criteria, resulted in 15 admissible articles.

Regarding the EEG recording parameters, although the electrode density did not directly affect the quality of the recorded signal, it did so to its spatial resolution, because a high density of electrodes increases the number of EEG signals and therefore the amount of information, favoring a more complete dataset that benefits the learning of the models. The number of electrodes considered was heterogeneous, although montages with 14 electrodes predominated, followed by 64 and 32 channels in the same proportion.

### 4.2. Data Pre-Processing

Table 3 showed that 27.91% of the studies did not specify the EEG cleaning process, 23.26% carried out little pre-processing in the data (which is the application of filters without eliminating artifacts), and 48.84% of the articles used artifact-free EEG signals. In



the case of those studies that did not specify the EEG cleaning, it was assumed that they did not perform any pre-processing of the signals. When the results of these three types of cleaning for the same model and the same objective problem were compared, it was observed that the articles [38,49,65] which corresponded to artifact-free signals, signals with unspecified cleaning, and signals with little preprocessing, respectively, used CNN in the classification problem of patients with PD vs. healthy controls, and obtained accuracy values of 100%, 99.62%, and 99.22%, and sensitivity values of 100%, 99.17% and 98.98% for the articles mentioned, respectively. It can be noted that the values obtained in both metrics are independent of the EEG cleaning procedure. This also occurs in the articles [51,66], which considered artifact-free signals and unspecified cleanliness, respectively, and used MLP in the same classification problem, obtaining sensitivity values of 100% and 96.7%, and specificity values of 94% and 100%, respectively. These results reinforce the conclusions that were obtained in [28], where it was indicated that the EEG cleaning protocol did not influence the results of the ML models.

Regarding the extracted characteristics, we found a great heterogeneity, although the use of spectral characteristics predominates in 88.37% of the articles. This may be due to the fact that, especially in PD, no visual alterations are observed in the EEG signals of PD patients, whereas a spectral analysis provides information on variations in the EEG bands, which are related to alterations in the patients' condition and are therefore applicable in the clinical setting.

#### 4.3. Models Evaluation

There are two sets of studies: those that compared different models and those that used a single model. The latter should provide a detailed bibliographic analysis that serves as a comparison and justified the development carried out and the results obtained. Delving into the evaluation of the models, it is worth noting that both the architecture of the models and the training/validation methods were not specified in all articles. Regarding the architecture, even if it was specified, the information provided for different models within the same study exhibited a different degree of specificity, turning the architecture into heterogeneous information, which made it difficult to establish conclusions from a comparison between the articles. These facts are quite remarkable, since as shown in [41], the parameters that define the model architecture can greatly influence the results of the precision metrics, so it is important to specify them to increase the quality of the study.

The validation method is key to evaluate the training process and to obtain information on the generalizability of the model, facilitating a possible application in the clinical setting. The most widely used validation method was k-fold cross-validation, considered by 55.81% of the articles. Among them, 13 studies (30.23% of the total set) used 10-fold cross-validation and five studies (11.63%) used a five-fold cross-validation. The articles [37,58,59] did not specify the validation process and therefore the results cannot be generalized. This validation method allows for the evaluation of the model by minimizing the bias produced by the choice of the data. Hence, it is especially appropriate in those cases in which datasets are considered small or difficult to expand, much like when dealing with clinical data. Even though most of the studies specified the validation method, which constituted a good practice, it should be noted that it was not specified for all models whether the results obtained belonged to the training, validation, or test sets, although the three provide crucial information about the training process of the model and its ability to generalize to blind data. This was an inconvenience when comparing the results of the models since the value of the metrics may be rather different in each of these scenarios.

Finally, it was striking that the accuracy metric was the most widespread metric among the selected articles, being present in 90.70% of the studies, followed by the sensitivity metric considered in 69.77%. It should be noted that in the most recent articles, the evaluation of the models was carried out by means of other metrics. In the studies that work with patients, especially in a problem of diagnosis of a disease, the sensitivity or recall, specificity, and precision metrics are also very useful to evaluate the results of the model, as they

provide a measure of the correctly classified positives, the correctly classified negatives, and the hits among the positives predictions, respectively.

#### 4.4. Global Discussion

A different set of papers focused on classifying PD patients or evaluating cognitive impairment. The articles related to the detection of walking alterations presented deficiencies in terms of the patient demographic data and medication status. Specifically, refs. [29–31,36,38,39,44,48,49,52,57,67] were the articles that considered classification problems and went through the quality criteria specified above. All of them considered balanced classes. Article [48] may also be discarded, as it did not consider balanced classes, so the results could be affected by the majority class. Among the remaining studies, article [36] was also excluded from the comparison, since its objective was the early diagnosis and the results obtained may differ greatly from the classification of patients in advanced stages of the disease. Given that the rest of the articles had accuracy as a common metric, this metric was evaluated for the problem of diagnosing PD by means of EEG through ML techniques, obtaining a value of  $97.35 \pm 3.46\%$ . Consequently, this value could be considered as a baseline for future studies that focus on this diagnostic problem. The sensitivity metric appeared in eight of the articles with a mean value of  $96.36 \pm 5.00\%$ . These studies were performed with different EEG channels, predominating the setups with 14 and 32 electrodes. The low variation in the value provided by the metrics may lead us to think that the number of channels did not influence the classification results.

Regarding the models used in recent years (from 2019 to 2021), and taking into account all the results in Table 3, it can be observed that neural models are increasingly used in the classification problem, appearing in 11 of the 16 studies within these years. However, there was no standard in the cleaning process. In 2019, there was an equitable distribution between the three types of processing. In 2020, three of the five articles considered did not perform EEG pre-processing, whereas in 2021, of the nine studies, five considered the removal of artifacts. This analysis indicates, again, that the cleaning protocol did not influence the results of the study. The use of increasingly complex techniques to solve the problem is striking, highlighting the use of CNNs and hybrid models based on CNN + RNN.

The articles [32,34,36,45,47,53] classified the level of cognitive impairment of PD patients. It should be noted that all of them have been published in the years 2019 and 2021, indicating the novelty of the subject and marking the evolution in the study and diagnosis of PD. In all of them, the elimination of artifacts in the EEG signals was carried out, and the use of the SVM and RF models stands out. Given the novelty of the research topic, it was expected that the models initially used were non-neuronal and that the pre-processing sought to minimize the noisy component from the EEG signals. For this case, we did not consider that there were enough studies that met the proposed quality criteria (description of the medication and patient status, use of balanced classes, and specification of the proposed validation method) to obtain a baseline for this case. However, the mean provided by those studies that passed the cut-off [45,47] was 77% accuracy. Although [47] did not use fully balanced classes, it did take these differences into account in evaluating the results and performed a five-sample multi-class task, whereas [45] used balanced classes in a binary problem.

In light of the previous analysis, growth in the complexity of the neural models, and in the number of studies that address the problem of identifying the cognitive status in patients with PD, is expected. This will benefit the development of a marker of the disease in the early stages. Regarding the methodology, it would be desirable that the studies provide more specific information about the state of the patients and their demographics. This would allow for an objective evaluation of the results obtained by the models, which should be based on a cross-validation process.

## 5. Conclusions

The objective of this review consisted in the study of ML techniques applied to the analysis of EEG associated with PD. The search process carried out on 14 February 2022 yielded 358 results that were submitted to the selection process following the PRISMA guidelines. This process resulted in 59 articles dealing with this topic, from different perspectives or with different objectives. Although they were mainly focused on the diagnosis of PD, studies on the classification of cognitive impairment and prediction of alterations in walking were also found. These studies were analyzed according to three key points in the development of ML techniques, which were the dataset quality, the data preprocessing, and the model used.

The most widely used models were SVM and ANN, with ANN encompassing both MLPs and those ANNs not specified (but suggested to be MLPs), followed by KNN, RF, and CNN. Although the most used individual model was the SVM, the DL techniques, with CNN models, predominated in the diagnostic problem in recent years. Currently, research has focused on the identification of cognitive impairment in patients with PD, in light of articles from recent years (2019 and 2021). The importance of the validation process should also be highlighted. In particular, the k-fold cross-validation method (with  $k = 10$ ) was used in most of the articles, which allowed the objective evaluation of the results, eliminating the bias produced by the choice of the subjects and facilitating future applications in the clinical field.

Among the main results provided by this analysis, it was found that neither the EEG cleaning protocol nor the number of channels were relevant to improve the performance in the classification or diagnosis of PD, although the 14-electrode setup was the most used in these studies. In addition, the content of the studies whose objective was the diagnosis of PD was evaluated. Within this group of articles, those satisfying quality criteria were selected to provide a baseline for the accuracy metric, yielding a value of  $97.35 \pm 3.46\%$ .

According to the analysis performed, future research directions are recommended to be based on the identification of the cognitive status of PD patients, to aid in the early diagnosis or staging of the disease. Moreover, studies are encouraged to incorporate and take into account patient demographics, carry out a k-fold CV for results validation, and use various metrics to provide a global view of the model's performance.

**Author Contributions:** Conceptualization, A.M.M., Á.J.G.-T. and J.P.R.M.; methodology, A.M.M., Á.J.G.-T. and J.P.R.M.; software, A.M.M.; validation, A.M.M., Á.J.G.-T. and J.P.R.M.; formal analysis, A.M.M.; investigation, A.M.M., Á.J.G.-T. and J.P.R.M.; resources, Á.J.G.-T. and J.P.R.M.; data curation, A.M.M.; writing—original draft preparation, A.M.M., Á.J.G.-T. and J.P.R.M.; writing—review and editing, A.M.M., Á.J.G.-T. and J.P.R.M.; visualization, A.M.M., Á.J.G.-T. and J.P.R.M.; supervision, Á.J.G.-T. and J.P.R.M.; project administration, Á.J.G.-T. and J.P.R.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** We thank Silvia Romero Azpitarte for her contribution to the screening process.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. National Institute of Neurological Disorders and Stroke. *Parkinson's Disease: Challenges, Progress, and Promise*; Publication No 15-5595; NIH: Bethesda, MD, USA, 2015. Available online: <https://www.ninds.nih.gov/health-information/patient-caregiver-education/hope-through-research/parkinsons-disease/parkinsons-disease-challenges-progress-and-promise> (accessed on 1 July 2022).
2. Dickson, D.W. Neuropathology of Parkinson disease. *Parkinsonism Relat. Disord.* **2018**, *46*, S30–S33. [CrossRef] [PubMed]

3. Jankovic, J. Progression of Parkinson Disease: Are We Making Progress in Charting the Course? *Arch. Neurol.* **2005**, *62*, 351–352. [[CrossRef](#)]
4. Beitz, J.M. Parkinson's Disease: A Review. *Front. Biosci.* **2014**, *6*, 65–74. [[CrossRef](#)] [[PubMed](#)]
5. Djaldetti, R.; Ziv, I.; Melamed, E. The mystery of motor asymmetry in Parkinson's disease. *Lancet Neurol.* **2006**, *5*, 796–802. [[CrossRef](#)]
6. Kostrzewa, R.M.; Nowak, P.; Kostrzewa, J.P.; Kostrzewa, R.A.; Brus, R. Peculiarities of L-DOPA treatment of Parkinson's disease. *Amino Acids* **2005**, *28*, 157–164. [[CrossRef](#)] [[PubMed](#)]
7. Kalia, L.V.; Lang, A.E. Parkinson's Disease. *Lancet* **2015**, *386*, 896–912. [[CrossRef](#)]
8. Beach, T.G.; Adler, C.H. Importance of low diagnostic Accuracy for early Parkinson's disease. *Mov. Disord. Off. J. Mov. Disord. Soc.* **2018**, *33*, 1551–1554. [[CrossRef](#)]
9. Gandar, M.J.; Edgar, J.C.; Klook, K.; Siegel, S.J. Gamma synchrony: Towards a translational biomarker for the treatment-resistant symptoms of schizophrenia. *Neuropharmacology* **2012**, *62*, 1504–1518. [[CrossRef](#)]
10. Smailovic, U.; Jelic, V. Neurophysiological Markers of Alzheimer's Disease: Quantitative EEG Approach. *Neurol. Ther.* **2019**, *8*, 37–55. [[CrossRef](#)]
11. Kannathal, N.; Choo, M.L.; Acharya, U.R.; Sadasivan, P.K. Entropies for detection of epilepsy in EEG. *Comput. Methods Programs Biomed.* **2005**, *80*, 187–194. [[CrossRef](#)]
12. Bigdely-Shamlo, N.; Mullen, T.; Kothe, C.; Su, K.-M.; Robbins, K.A. The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Front. Neuroinform.* **2015**, *9*, 16. [[CrossRef](#)] [[PubMed](#)]
13. Cole, S.; Voytek, B. Cycle-by-cycle analysis of neural oscillations. *J. Neurophysiol.* **2019**, *122*, 849–861. [[CrossRef](#)] [[PubMed](#)]
14. Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. II-Recent Progress. In *Computer Games I*; Levi, D.N.L., Ed.; Springer: New York, NY, USA, 1988; pp. 366–400. [[CrossRef](#)]
15. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciampi, F.; Ghafoorian, M.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)] [[PubMed](#)]
16. Jiang, F.; Jiang, Y.; Zhi, H.; Dong, Y.; Li, H.; Ma, S.; Wang, Y. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc. Neurol.* **2017**, *2*, 230–243. [[CrossRef](#)] [[PubMed](#)]
17. Miller, D.D.; Brown, E.W. Artificial Intelligence in Medical Practice: The Question to the Answer? *Am. J. Med.* **2018**, *131*, 129–133. [[CrossRef](#)] [[PubMed](#)]
18. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
19. Ru, G.; Crescio, M.I.; Ingravalle, F.; Maurella, C.; Gregori, D.; Lanera, C.; Azzolina, D.; Lorenzoni, G.; Soriani, N.; Zec, S.; et al. Machine Learning Techniques applied in risk assessment related to food safety. *EFSA Supporting Publ.* **2017**, *14*, EN-1254. [[CrossRef](#)]
20. Roy, Y.; Banville, H.; Albuquerque, I.; Gramfort, A.; Falk, T.H.; Faubert, J. Deep learning-based electroencephalography analysis: A systematic review. *J. Neural Eng.* **2019**, *16*, 051001. [[CrossRef](#)]
21. Sheng, J.; Wang, B.; Zhang, Q.; Liu, Q.; Ma, Y.; Liu, W.; Shao, M.; Chen, B. A novel joint HCPMMP method for automatically classifying Alzheimer's and different stage MCI patients. *Behav. Brain Res.* **2019**, *365*, 210–221. [[CrossRef](#)]
22. Raghavendra, U.; Acharya, U.R.; Adeli, H. Artificial Intelligence Techniques for Automated Diagnosis of Neurological Disorders. *Eur. Neurol.* **2019**, *82*, 41–64. [[CrossRef](#)]
23. Jahmunah, V.; Oh, S.L.; Rajinikanth, V.; Ciaccio, E.J.; Cheong, K.H.; Arunkumar, N.; Acharya, U.R. Automated detection of schizophrenia using nonlinear signal processing methods. *Artif. Intell. Med.* **2019**, *100*, 101698. [[CrossRef](#)] [[PubMed](#)]
24. Chen, Y.; Gong, C.; Hao, H.; Guo, Y.; Xu, S.; Zhang, Y.; Yin, G.; Cao, X.; Yang, A.; Meng, F.; et al. Automatic Sleep Stage Classification Based on Subthalamic Local Field Potentials. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 118–128. [[CrossRef](#)] [[PubMed](#)]
25. Zhou, M.; Tian, C.; Cao, R.; Wang, B.; Niu, Y.; Hu, T.; Guo, H.; Xiang, J. Epileptic Seizure Detection Based on EEG Signals and CNN. *Front. Neuroinform.* **2018**, *12*, 95. [[CrossRef](#)] [[PubMed](#)]
26. Dhivya, S.; Nithya, A. A Review on Machine Learning Algorithm for EEG Signal Analysis. In Proceedings of the Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 29–31 March 2018; pp. 54–57. [[CrossRef](#)]
27. Rasheed, K.; Qayyum, A.; Qadir, J.; Sivathamboo, S.; Kwan, P.; Kuhlmann, L.; O'Brien, T.; Razi, A. Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review. *arXiv* **2020**, arXiv:2002.01925. [[CrossRef](#)] [[PubMed](#)]
28. Maitin, A.M.; García-Tejedor, A.J.; Romero Muñoz, J.P. Machine Learning Approaches for Detecting Parkinson's Disease from EEG Analysis: A Systematic Review. *Appl. Sci.* **2020**, *10*, 8662. [[CrossRef](#)]
29. Lee, S.; Hussein, R.; McKeown, M.J. A deep convolutional-recurrent neural network architecture for Parkinson's disease EEG classification. In Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP), Ottawa, ON, Canada, 11–14 November 2019; pp. 1–4. [[CrossRef](#)]
30. Lee, S.; Hussein, R.; Ward, R.; Wang, A.J.; McKeown, M.J. A convolutional-recurrent neural network approach to resting-state EEG classification in Parkinson's disease. *J. Neurosci. Methods* **2021**, *361*, 109282. [[CrossRef](#)]
31. Oh, S.L.; Hagiwara, Y.; Raghavendra, U.; Yuvaraj, R.; Arunkumar, N.; Murugappan, M.; Acharya, U.R. A deep learning approach for Parkinson's disease diagnosis from EEG signals. *Neural Comput. Appl.* **2020**, *32*, 10927–10933. [[CrossRef](#)]



32. Chaturvedi, M.; Bogaarts, J.G.; Kozak, V.V.; Hatz, F.; Gschwandtner, U.; Meyer, A.; Fuhr, P.; Roth, V. Phase lag index and spectral power as QEEG features for identification of patients with mild cognitive impairment in Parkinson's disease. *Clin. Neurophysiol.* **2019**, *130*, 1937–1944. [[CrossRef](#)]
33. Chaturvedi, M.; Hatz, F.; Gschwandtner, U.; Bogaarts, J.G.; Meyer, A.; Fuhr, P.; Roth, V. Quantitative EEG (QEEG) measures differentiate Parkinson's disease (PD) patients from healthy controls (HC). *Front. Aging Neurosci.* **2017**, *9*, 3. [[CrossRef](#)]
34. Geraedts, V.J.; Koch, M.; Contarino, M.F.; Middelkoop, H.A.M.; Wang, H.; van Hilten, J.J.; Back, T.H.W.; Tannemaat, M.R. Machine learning for automated EEG-based biomarkers of cognitive impairment during Deep Brain Stimulation screening in patients with Parkinson's Disease. *Clin. Neurophysiol.* **2021**, *132*, 1041–1048. [[CrossRef](#)]
35. Handojoseno, A.M.A.; Naik, G.R.; Gilat, M.; Shine, J.M.; Nguyen, T.N.; Ly, Q.T.; Lewis, S.J.G.; Nguyen, H.T. Prediction of Freezing of Gait in Patients with Parkinson's Disease Using EEG Signals. *Stud. Health Technol. Inform.* **2018**, *246*, 124–131. [[PubMed](#)]
36. Hassin-Baer, S.; Cohen, O.S.; Israeli-Korn, S.; Yahalom, G.; Benizri, S.; Sand, D.; Issachar, G.; Geva, A.B.; Shani-Hershkovich, R.; Peremen, Z. Identification of an early-stage Parkinson's disease neuromarker using event-related potentials, brain network analytics and machine-learning. *PLoS ONE* **2022**, *17*, e0261947. [[CrossRef](#)]
37. Kamalraj, S.; Rejith, K.N.; Prasanna Venkatesan, G.K.D. Frequency domain analysis for the classification of Parkinson's disease patients. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Tamil Nadu, India, 12–13 April 2019; Volume 561. [[CrossRef](#)]
38. Khare, S.K.; Bajaj, V.; Acharya, U.R. PDCNNNet: An Automatic Framework for the Detection of Parkinson's Disease Using EEG Signals. *IEEE Sens. J.* **2021**, *21*, 17017–17024. [[CrossRef](#)]
39. Loh, H.W.; Ooi, C.P.; Palmer, E.; Barua, P.D.; Dogan, S.; Tuncer, T.; Baygin, M.; Acharya, U.R. GaborPDNet: Gabor Transformation and Deep Neural Network for Parkinson's Disease Detection Using EEG Signals. *Electronics* **2021**, *10*, 1740. [[CrossRef](#)]
40. Ly, Q.T.; Handojoseno, A.M.A.; Gilat, M.; Nguyen, N.; Chai, R.; Tran, Y.; Lewis, S.J.G.; Nguyen, H.T. Identifying montages that best detect the electroencephalogram power spectrum alteration during freezing of gait in Parkinson's disease patients. In Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 6094–6097. [[CrossRef](#)]
41. Saikia, A.; Hussain, M.; Barua, A.R.; Paul, S. Performance analysis of various neural network functions for Parkinson's disease classification using EEG and EMG. *Int. J. Innov. Technol. Explor. Eng.* **2019**, *9*, 3402–3406. [[CrossRef](#)]
42. Shi, X.; Wang, T.; Wang, L.; Liu, H.; Yan, N. Hybrid convolutional recurrent neural networks outperform CNN and RNN in Task-state EEG detection for Parkinson's disease. In Proceedings of the 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Lanzhou, China, 18–21 November 2019; pp. 939–944. [[CrossRef](#)]
43. Vanegas, M.I.; Ghilardi, M.F.; Kelly, S.P.; Blangero, A. Machine learning for EEG-based biomarkers in Parkinson's disease. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 3–6 December 2018; pp. 2661–2665. [[CrossRef](#)]
44. Waninger, S.; Berka, C.; Karic, M.S.; Korszen, S.; Mozley, P.D.; Henschcliffe, C.; Kang, Y.; Hesterman, J.; Mangoubi, T.; Verma, A. Neurophysiological Biomarkers of Parkinson's Disease. *J. Parkinson's Dis.* **2020**, *10*, 471–480. [[CrossRef](#)]
45. Zhang, J.; Gao, Y.; He, X.; Feng, S.; Hu, J.; Zhang, Q.; Zhao, J.; Huang, Z.; Wang, L.; Ma, G.; et al. Identifying Parkinson's disease with mild cognitive impairment by using combined MR imaging and electroencephalogram. *Eur. Radiol.* **2021**, *31*, 7386–7394. [[CrossRef](#)]
46. Anwar, T.; Rehmat, N.; Naveed, H. A Generic Approach for Classification of Psychological Disorders Diagnosis using EEG. In Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Guadalajara, Jalisco, Mexico, 1–5 November 2021; pp. 2025–2029. [[CrossRef](#)]
47. Betrouni, N.; Delval, A.; Chaton, L.; Defebvre, L.; Duits, A.; Moonen, A.; Leentjens, A.F.G.; Dujardin, K. Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: Preliminary results. *Mov. Disord.* **2019**, *34*, 210–217. [[CrossRef](#)]
48. Chu, C.; Zhang, Z.; Wang, J.; Liu, S.; Wang, F.; Sun, Y.; Han, X.; Li, Z.; Zhu, X.; Liu, C. Deep learning reveals personalized spatial spectral abnormalities of high delta and low alpha bands in EEG of patients with early Parkinson's disease. *J. Neural Eng.* **2021**, *18*, 066036. [[CrossRef](#)]
49. Emamzadeh-Hashemi, E.A.; Mahdizadeh, A.; Mirian, M.S.; Lee, S.; McKeown, M.J. Deep transfer learning for parkinson's disease monitoring by image-based representation of resting-state EEG using directional connectivity. *Algorithms* **2022**, *15*, 5. [[CrossRef](#)]
50. Guo, G.; Wang, S.; Wang, S.; Zhou, Z.; Pei, G.; Yan, T. Diagnosing Parkinson's Disease Using Multimodal Physiological Signals. In *Human Brain and Artificial Intelligence. HBAI 2021. Communications in Computer and Information Science, Yokohama, Japan, January 7, 2021*; Wang, Y., Ed.; Springer: Singapore, 2021; Volume 1369. [[CrossRef](#)]
51. Jervis, B.W.; Saatchi, M.R.; Lacey, A.; Roberts, T.; Allen, E.M.; Hudson, N.R.; Oke, S.; Grimsley, M. Artificial neural network and spectrum analysis methods for detecting brain diseases from the CNV response in the electroencephalogram. *IEE Proc. Sci. Meas. Technol.* **1994**, *141*, 432–440. [[CrossRef](#)]
52. Khare, S.K.; Bajaj, V.; Acharya, U.R. Detection of Parkinson's disease using automated tunable Q wavelet transform technique with EEG signals. *Biocybern. Biomed. Eng.* **2021**, *41*, 679–689. [[CrossRef](#)]
53. Koch, M.; Geraedts, V.; Wang, H.; Tannemaat, M.; Back, T. Automated Machine Learning for EEG-Based Classification of Parkinson's Disease Patients. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 4845–4852. [[CrossRef](#)]

54. Liu, S.; Li, M.; Feng, Y.; Zhang, M.; Acquah, M.E.E.; Huang, S.; Chen, J.; Ren, P. Brain Network Analysis by Stable and Unstable EEG Components. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 1080–1092. [[CrossRef](#)] [[PubMed](#)]
55. Ly, Q.T.; Gilat, M.; Chai, R.; Martens, K.A.E.; Georgiades, M.; Naik, G.R.; Tran, Y.; Lewis, S.J.G.; Nguyen, H.T. Detection of turning freeze in Parkinson's disease based on S-transform decomposition of EEG signals. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju Island, Korea, 11–15 July 2017; pp. 3044–3047. [[CrossRef](#)]
56. Ly, Q.T.; Handojoseno, A.M.A.; Gilat, M.; Nguyen, N.; Chai, R.; Tran, Y.; Lewis, S.J.G.; Nguyen, H.T. Detection of Gait Initiation Failure in Parkinson's disease patients using EEG signals. In Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 1599–1602. [[CrossRef](#)]
57. Oliveira, A.P.S.; de Santana, M.A.; Andrade, M.K.S.; Gomes, J.C.; Rodrigues, M.C.A.; dos Santos, W.P. Early diagnosis of Parkinson's disease using EEG, machine learning and partial directed coherence. *Res. Biomed. Eng.* **2020**, *36*, 311–331. [[CrossRef](#)]
58. Rejith, K.N.; Subramaniam, K. Analysis of emotional states in Parkinson's disease using entropy, energy-entropy and teager energy-entropy features. *Indian J. Public Health Res. Dev.* **2018**, *9*, 1099–1102. [[CrossRef](#)]
59. Rejith, K.N.; Subramaniam, K. Classification of emotional states in Parkinson's disease patients using machine learning algorithms. *Biomed. Pharmacol. J.* **2018**, *11*, 333–341. [[CrossRef](#)]
60. Rodrigues, P.M.; Teixeira, J.P. Classification of electroencephalogram signals using artificial neural networks. In Proceedings of the 2010 3rd International Conference on Biomedical Engineering and Informatics, Yantai, China, 16–18 October 2010; pp. 808–812. [[CrossRef](#)]
61. Ruffini, G.; Ibañez, D.; Castellano, M.; Dubreuil-Vall, L.; Soria-Frisch, A.; Postuma, R.; Gagnon, J.-F.; Montplaisir, J. Deep learning with EEG spectrograms in rapid eye movement behavior disorder. *Front. Neurol.* **2019**, *10*, 806. [[CrossRef](#)]
62. Ruffini, G.; Ibañez, D.; Castellano, M.; Dunne, S.; Soria-Frisch, A. EEG-driven RNN classification for prognosis of neurodegeneration in at-risk patients. *Lect. Notes Comput. Sci. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinform.)* **2016**, *9886*, 306–313. [[CrossRef](#)]
63. Saikia, A.; Hussain, M.; Barua, A.R.; Paul, S. EEG-EMG correlation for Parkinson's disease. *Int. J. Eng. Adv. Technol.* **2019**, *8*, 1179–1185. [[CrossRef](#)]
64. Shaban, M. Automated Screening of Parkinson's Disease Using Deep Learning Based Electroencephalography. In Proceedings of the 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER), Virtually, 4–6 May 2021; pp. 158–161. [[CrossRef](#)]
65. 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](#)]
66. Shreya Prabhu, K.; Martis, R.J. Diagnosis of Parkinson's Disease using Computer Aided Tool based on EEG. In Proceedings of the IEEE 17th India Council International Conference INDICON, New Delhi, India, 10–13 December 2020; pp. 1–4. [[CrossRef](#)]
67. Yuvaraj, R.; Acharya, U.R.; Hagiwara, Y. A novel Parkinson's Disease Diagnosis Index using higher-order spectra features in EEG signals. *Neural Comput. Appl.* **2018**, *30*, 1225–1235. [[CrossRef](#)]
68. Handojoseno, A.M.A.; Shine, J.M.; Nguyen, T.N.; Tran, Y.; Lewis, S.J.G.; Nguyen, H.T. The Detection of Freezing of Gait in Parkinson's Disease Patients Using EEG Signals Based on Wavelet Decomposition. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, August 28–September 1 2012; pp. 69–72. [[CrossRef](#)]
69. Handojoseno, A.M.A.; Shine, J.M.; Nguyen, T.N.; Tran, Y.; Lewis, S.J.G.; Nguyen, H.T. Using EEG Spatial Correlation, Cross Frequency Energy, and Wavelet Coefficients for the Prediction of Freezing of Gait in Parkinson's Disease Patients. In Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; pp. 4263–4266. [[CrossRef](#)]
70. Rahman, M.A.; Tutul, A.A.; Islam, A.B.M.A.A. Solving the Maze of Diagnosing Parkinson's Disease based on Portable EEG Sensing to be Adaptable to Go In-The-Wild. In Proceedings of the 7th International Conference on Networking, Systems and Security, Dhaka, Bangladesh, 22–24 December 2020; pp. 65–73. [[CrossRef](#)]
71. Vanneste, S.; Song, J.-J.; Ridder, D.D. Thalamocortical Dysrhythmia Detected by Machine Learning. *Nat. Commun.* **2018**, *9*, 1103. [[CrossRef](#)] [[PubMed](#)]
72. Yuvaraj, R.; Murugappan, M.; Ibrahim, N.M.; Sundaraj, K.; Omar, M.I.; Mohamad, K.; Palaniappan, R. Optimal set of EEG features for emotional state classification and trajectory visualization in Parkinson's disease. *Int. J. Psychophysiol.* **2014**, *94*, 482–495. [[CrossRef](#)] [[PubMed](#)]
73. Yuvaraj, R.; Murugappan, M.; Ibrahim, N.M.; Omar, M.I.; Sundaraj, K.; Mohamad, K.; Palaniappan, R.; Satiyan, M. Emotion classification in Parkinson's disease by higher-order spectra and power spectrum features using EEG signals: A comparative study. *J. Integr. Neurosci.* **2014**, *13*, 89–120. [[CrossRef](#)]
74. Yuvaraj, R.; Murugappan, M.; Ibrahim, N.M.; Sundaraj, K.; Omar, M.I.; Mohamad, K.; Palaniappan, R. Detection of emotions in Parkinson's disease using higher order spectral features from brain's electrical activity. *Biomed. Signal Processing Control.* **2014**, *14*, 108–116. [[CrossRef](#)]
75. Murugappan, M.; Alshuaib, W.B.; Bourisly, A.; Sruthi, S.; Ranjana, R. Recurrence Quantification Analysis based Emotion Detection in Parkinson's disease using EEG Signals. In Proceedings of the 4th International Conference on Computer, Communication and Signal Processing (ICCCSP), Chennai, India, 28–29 April 2020; pp. 1–6. [[CrossRef](#)]



76. Murugappan, M.; Alshuaib, W.; Bourisly, A.K.; Khare, S.K.; Sruthi, S.; Bajaj, V. Tunable Q wavelet transform based emotion classification in Parkinson's disease using Electroencephalography. *PLoS ONE* **2020**, *15*, e0242014. [[CrossRef](#)]
77. Cesari, M.; Christensen, J.A.E.; Muntean, M.L.; Mollenhauer, B.; Sixel-Döring, F.; Sorensen, H.B.D.; Trenkwalder, C.; Jennum, P. A data-driven system to identify REM sleep behavior disorder and to predict its progression from the prodromal stage in Parkinson's disease. *Sleep Med.* **2020**, *77*, 238–248. [[CrossRef](#)]
78. Sorensen, G.L.; Jennum, P.; Kempfner, J.; Zoetmulder, M.; Sorensen, H.B.D. A Computerized Algorithm for Arousal Detection in Healthy Adults and Patients with Parkinson Disease. *J. Clin. Neurophysiol.* **2012**, *29*, 58–64. [[CrossRef](#)]
79. Patanaik, A.; Ong, J.L.; Gooley, J.J.; Ancoli-Israel, S.; Chee, M.W.L. An end-to-end framework for real-time automatic sleep stage classification. *Sleep* **2018**, *41*, zsy041. [[CrossRef](#)]
80. Sorensen, G.L.; Kempfner, J.; Jennum, P.; Sorensen, H.B.D. Detection of Arousals in Parkinson's Disease Patients. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; pp. 2764–2767. [[CrossRef](#)]
81. Castano-Candamil, S.; Piroth, T.; Reinacher, P.; Sajonz, B.; Coenen, V.A.; Tangermann, M. Identifying controllable cortical neural markers with machine learning for adaptive deep brain stimulation in Parkinson's disease. *Neuroimage Clin.* **2020**, *28*, 102376. [[CrossRef](#)]
82. Geraedts, V.J.; Koch, M.; Kuiper, R.; Kefalas, M.; Back, T.H.W.; van Hilten, J.J.; Wang, H.; Middelkoop, H.A.; Gaag, N.A.; Contarino, M.F.; et al. Preoperative Electroencephalography-Based Machine Learning Predicts Cognitive Deterioration after Subthalamic Deep Brain Stimulation. *Mov. Disord.* **2021**, *36*, 2324–2334. [[CrossRef](#)] [[PubMed](#)]
83. Stuart, M.; Wickramasinghe, C.S.; Marino, D.L.; Kumbhare, D.; Holloway, K.; Manic, M. Machine Learning for Deep Brain Stimulation Efficacy using Dense Array EEG. In Proceedings of the 2019 12th International Conference on Human System Interaction (HSI), Richmond, VA, USA, 25–27 June 2019; pp. 143–150. [[CrossRef](#)]
84. Sand, D.; Arkadir, D.; Snineh, M.A.; Marmor, O.; Israel, Z.; Bergman, H.; Hassin-Baer, S.; Israeli-Korn, S.; Peremen, Z.; Geva, A.B.; et al. Deep Brain Stimulation Can Differentiate Subregions of the Human Subthalamic Nucleus Area by EEG Biomarkers. *Front. Syst. Neurosci.* **2021**, *15*, 747681. [[CrossRef](#)] [[PubMed](#)]
85. Maurer, A.; Hanrahan, S.; Nedrud, J.; Hebb, A.O.; Papandreou-Suppappola, A. Suppression of Neurostimulation Artifacts and Adaptive Clustering of Parkinson's Patients Behavioral Tasks using EEG. In Proceedings of the 50th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 6–9 November 2016; pp. 851–855. [[CrossRef](#)]
86. Saikia, A.; Majhi, V.; Hussain, M.; Barua, A.R.; Paul, S.; Verma, J.K. Machine Learning based Diagnostic System for Early Detection of Parkinson's Disease. In Proceedings of the International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2–4 July 2020; pp. 275–279. [[CrossRef](#)]
87. Geman, O.; Chiuchisan, I.; Covasa, M.; Eftaxias, K.; Sanei, S.; Ferreira Madeira, J.G.; Mancebo Boley, R.A. Joint EEG-EMG Signal Processing for Identification of the Mental Tasks in Patients with Neurological Diseases. In Proceedings of the 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 29 August–2 September 2016; pp. 1598–1602. [[CrossRef](#)]
88. Barrachina-Fernández, M.; Maitín, A.M.; Sánchez-Ávila, C.; Romero, J.P. Wearable Technology to Detect Motor Fluctuations in Parkinson's Disease Patients: Current State and Challenges. *Sensors* **2021**, *21*, 4188. [[CrossRef](#)] [[PubMed](#)]