Expert Systems With Applications 212 (2023) 118772

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Parkinson's disease effective biomarkers based on Hjorth features improved by machine learning

Bruno Fonseca Oliveira Coelho [a,*], Ana Beatriz Rodrigues Massaranduba [a], Carolline Angela dos Santos Souza [b], Giovanni Guimarães Viana [b], Ivani Brys [a,c], Rodrigo Pereira Ramos [a,b]

[a] Postgraduate Program in Health and Biological Sciences, Federal University of Vale do São Francisco (UNIVASF), Petrolina, Brazil
[b] Electrical Engineering Faculty, Federal University of Vale do São Francisco (UNIVASF), Juazeiro, Bahia, Brazil
[c] Psychology Faculty, Federal University of Vale do São Francisco (UNIVASF), Petrolina, Brazil

ARTICLE INFO

Keywords:
Parkinson's disease
EEG
Hjorth features
Machine learning

ABSTRACT

Parkinson's disease (PD) is the second most common neurodegenerative condition in the world and is caused by reduced levels of dopamine in the central nervous system. The diagnosis of PD is a difficult and time-consuming task, and there is no definitive protocol for achieving it. Therefore, several studies have been performed in order to find reliable PD biomarkers. The analysis of characteristics of electroencephalogram (EEG) signals is one of the techniques that have been used in the search for biomarkers. EEG signals capture the activity of neurons through electrodes placed on the scalp and with the advancement of Artificial Intelligence (AI) techniques, their characteristics have started to be used in machine learning (ML) algorithms for the automatic diagnosis of brain diseases, suggesting that EEG signals are promising biomarkers that could be used for automatic identification of individuals with PD. Thus, this work evaluates the performance of Hjorth features obtained from electroencephalographic signals, as biomarkers for Parkinson's disease. Using the database available at the public repository called The Patient Repository for EEG Data + Computational Tools (PRED + CT), we analyzed EEG data from PD individuals periodically exposed to auditory stimuli. The analysis of the proposed biomarkers showed differences between healthy and PD patients in parietal, frontal, central, and occipital lobes. For classification the Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest algorithms were used followed by a 5-fold cross-validation methodology. The proposed model achieved an accuracy of 89.56% when differentiating patients with PD and healthy individuals with an SVM classifier. The results suggest that the Hjorth features extracted from EEG signals could be used as PD biomarkers.

1. Introduction

Parkinson's disease (PD) is a neurodegenerative pathology characterized by the loss of dopaminergic neurons in the substantia nigra and the accumulation of the protein α-synuclein (Poewe et al., 2017). Neurons are nervous system cells responsible for the transmission of nerve impulses. Unlike other cells of the human body, such as skin and bone cells, neurons are unable to repair themselves. Thus, with aging, these cells die and are never restored (Poewe et al., 2017). Due to this non-replacement of dead neurons, PD is considered a neurodegenerative disease that affects the central nervous system, causing the partial or total loss of motor functions, speech difficulties, dementia, amnesia, among other symptoms (Jankovic, 2008). Those are moderate symptoms in the initial stage and becomes more severe as it evolves with time.

PD diagnosis is a difficult and time-consuming task, and there is no definitive test for its determination. The lack of apparent motor or non-motor symptoms poses even more complexity on PD diagnosis, resulting in a high misdiagnosing rate (Barcelon et al., 2019). Some methodologies based on computer-aided diagnosis using artificial intelligence (AI) algorithms have been applied as a powerful tool for assisting professionals in the diagnosis of PD (Bhat et al., 2018).

Within this context, parameters extracted from voice recordings, gait patterns, and other signals have been proposed as biomarkers for Parkinson's disease. Voice signals are used due to vocal disorders

* Corresponding author.
E-mail addresses: brunofonsecaoc@gmail.com (B.F.O. Coelho), biamassaranduba@hotmail.com (A.B.R. Massaranduba), carolladss@gmail.com (C.A.d.S. Souza), giovanni.viana30@gmail.com (G.G. Viana), ivani.brys@univasf.edu.br (I. Brys), rodrigo.ramos@univasf.edu.br (R.P. Ramos).

https://doi.org/10.1016/j.eswa.2022.118772
Received 29 October 2021; Received in revised form 3 August 2022; Accepted 3 September 2022
Available online 14 September 2022

caused by the disease. Thus, patients with PD have difficulties in pronouncing vocal sounds, leaving speech markings that can be used for an early diagnosis of the disease (Sakar et al., 2013). Gait characteristics have also been considered as biomarkers for Parkinson's disease (Shetty & Rao, 2016; Wahid et al., 2015). Due to the motor limitations caused by neuronal degeneration, individuals with PD have a characteristic way of walking, which is used to identify people who suffer from this disease.

Electroencephalography (EEG) signals are quantitative biomarkers that have been recently used to predict PD. EEG signals measure amplitude variation or decay of the spectral power (Cavanagh et al., 2018; Klassen et al., 2011; Vanegas et al., 2019), these variations being correlated with the occurrence of the PD (Moisello et al., 2015). The EEG acquisition is performed by positioning electrodes along the scalp and measuring their voltage variations (Picard et al., 2001).

Although EEG signals have been used as possible PD biomarkers, some limitations must be observed, such as low spatial resolution, low signal-to-noise ratio, non-stationary signal, spurious elements, among others (Bigdely-Shamlo et al., 2015). Because of these problems, advanced signal processing techniques have been employed. As an example, one can cite the blind source separation techniques used for eliminating artifacts, such as those based on independent component analysis (ICA) or joint blind source separation (JBSS) (Chen et al., 2016). Characteristic extraction techniques are also important for the analysis and can be based on traditional techniques, such as the short-time Fourier transform (STFT) or the wavelet transform, as well as those based on high order statistics, which are more efficient as pattern discriminators, such as High Order Crossing (HOC) (Petrantonakis & Hadjileontiadis, 2009) and Hjorth features (Hjorth, 1970).

This work proposes a methodology to extract prominent features of EEG signals that can be used as biomarkers to identify patients with Parkinson's disease with the help of machine learning algorithms. The EEG data used in this work is part of the public repository The Patient Repository for EEG Data + Computational Tools (PRED+CT) (Cavanagh et al., 2017). Electroencephalographic signals were recorded while the participants were undergoing auditory stimuli. The stimuli were induced through the experiment known as the *3-Auditory Oddball Task*, and has been a methodology used by psychologists to study cognitive changes through event-related potentials on EEG signals (Potter et al., 2001). The biomarkers were obtained through the extraction of Hjorth features from the collected EEG signals. These features establish parameters that describe EEG signals in terms of amplitude and time, but still carry spectral information. Hjorth features were successfully employed for automatic recognition of emotions (Jenke et al., 2014; Patil et al., 2016) and for tremor identification in PD patients (Yao et al., 2018). After extracting the features as biomarkers, an intelligent computational algorithm was used to perform PD classification. Supervised machine learning algorithms based on Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Random Forest classifiers were used.

This paper is subsequently structured as follows. In Section 2, a list of works is presented which are related to this paper. Section 3 shows the methodology employed. In Section 4 the results and the corresponding discussion are presented. Finally, Section 5 states the conclusion of this work.

2. Related works

The use of Hjorth features to extract characteristics from EEG signals has been extensively studied in the literature. Jenke et al. (2014) used Hjorth features to recognize emotions from EEG signals. The application of machine learning algorithms in conjunction with Hjorth features to identify tremors in Parkinson patients was considered in Yao et al. (2018). Hjorth features extracted from local field potential signals in an adaptive deep brain stimulation system to predict motor impairment in PD were studied in Wang et al. (2018). These features were also used in a motor imagery system based on an EEG brain–computer interface (Oh et al., 2014).

The combined application of EEG signals and AI algorithms to identify Parkinson's conditions has also been previously investigated. Oh et al. (2018) analyzed the performance of a 13-layer convolutional neural network that was pre-trained to automatically detect seizures and cases of depression from EEG signals for PD identification. Koch et al. (2019) extracted statistical characteristics and power information in different frequency bands from the EEG signals acquired from 40 volunteers, which were used with other features, such as age and gender, in a Random Forest classifier.

EEG data collected during a visual stimulus experiment was considered in Vanegas et al. (2018), in which the participants visualized a static image containing only a gray background. Periodically, circles were added to this background, which blinked at a frequency of 25 Hz and had a duration of 2.4 s. The spectral amplitudes of the collected signals, computed through the Fast Fourier Transform (FFT), were applied as inputs for the classification algorithm based on decision trees to identify PD patients. Vanegas et al. (2019) also found that EEG frequency spectra was elevated in PD patients relative to control using a similar experiment. In de Oliveira et al. (2020), the authors also used EEG signals acquired during a visual task to evaluate the performance of nine classifiers to recognize PD patients. These classifiers had the partially directed coherence values computed between different electrode pairs as inputs. It is worth noting that only 10 electrodes were used, corresponding to the areas of the brain responsible for the propagation of visual information.

Yuvaraj et al. (2018) explored the use of Higher Order Spectra (HOS) as biomarkers for PD. EEG signals bispectra from resting patients were computed. Thirteen characteristics were evaluated based on the obtained bispectrum. SVM-type, K-nearest neighbors, and Naive Bayes classifiers were used in the proposed methodology.

Cavanagh et al. (2018), who also used the same database used in this article, employed an SVM classifier to differentiate PD from control subjects based on the spectral FFT coefficients over the trials. The present work proposes a different approach, in which the spectral FFT information is exchanged with Hjorth features, which showed to be reliable biomarker with improved performance.

3. Methodology

This section describes the data acquisition and further procedures employed for the evaluation of the features used for the classification of PD patients. The flowchart presented in Fig. 1 summarizes the methodology proposed in this work and will be detailed in the next sections.

3.1. Database and signals acquisition

The database used in this work is part of the public PRED+CT repository (Cavanagh et al., 2017). This repository hosts EEG data and analysis tools for the study of brain diseases. Data of 25 participants clinically diagnosed with Parkinson's disease and a matched control group consisting of an equivalent number of healthy individuals of corresponding gender and age. The participants underwent a task called Three-Stimulus Auditory Oddball Task, detailed in the next section. The group of patients diagnosed with PD participated in two sessions: in one of them, the individuals took their dopaminergic medicine normally; in the other, participants were instructed to not take any dopaminergic medicine in the last 15 h before the experiment. All participants were submitted to the Mini-Mental State Examination (MMSE). PD patients and control subjects had no significant differences in educational level, and no severe cases of dementia were identified through the test.

PD patients were evaluated in the Unified Parkinson's Disease Rating Scale (UPDRS) in periods on and off dopaminergic medication.

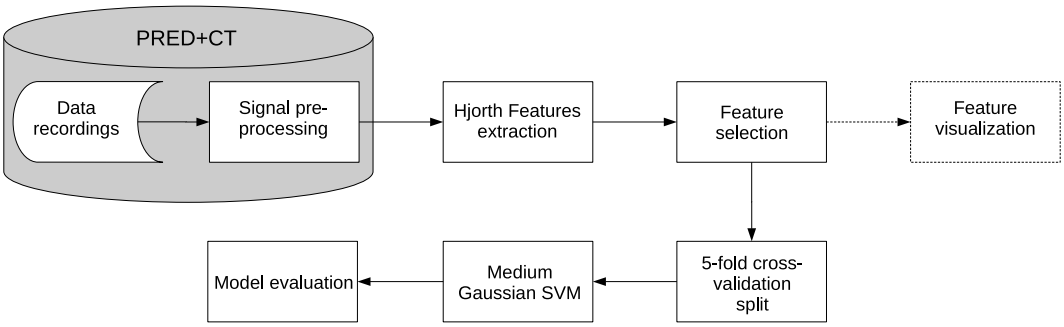

Fig. 1. Flowchart of the proposed methodology for PD identification.

The MMSE and UPDRS results, the levodopa dosage, the disease duration, and the patients' demographic information are provided in the supplementary file, adapted from Cavanagh et al. (2017).

EEG signals were continuously recorded through 64 AgCl electrodes with a reference in the AFz channel, a sampling rate of 500 Hz, filtered by a band pass filter (0.1 to 100 Hz). Signal pre-processing was performed on MATLAB and EEGLab toolbox following Cavanagh et al. (2018). Unreliable channels were removed, and only 60 channels were left. The signals were separated into epochs, with each epoch containing the 2 s preceding and 2 s following the sound stimulus. Bad channels and bad epochs were identified and subsequently removed. Signal artifacts such as ocular movement and eye blink were removed using independent component analysis (ICA). Finally, a re-referencing based on the average of all channels was performed. Data from one PD patient was not available in the database. Therefore, the signal from its corresponding control subject was also discarded. Thus, data from 48 participants were included in the present study (24 PD patients and 24 control subjects).

3.1.1. Three-stimulus auditory oddball task

The oddball paradigm is a commonly used task to assess cognition and attention with the help of EEG signals (Fichtenholtz et al., 2004; Potter et al., 2001). During the experiment, participants are periodically presented with two types of stimuli, named standard stimulus and target stimulus. In general, in the presence of a target stimulus, the participant must count (mentally or pressing a button) the occurrence of that stimulus, while the standard stimuli must not be counted.

The experiment employed a variation of the oddball paradigm called three-stimulus auditory oddball. It is composed of 3 types of auditory stimuli: pattern, target, and a divergent stimulus (Cavanagh et al., 2018). In this case, the stimuli consisted of three different sounds, and only the target signals should be counted by the participants. The standard signal (440 Hz tone) had an occurrence rate of 70%, the target signal (660 Hz tone) had an occurrence of 15%, and the divergent signal, represented by a naturalistic sound, had an occurrence of 15%. Two blocks of 100 stimuli each were performed. At the end of the experiment, the participants were asked to report the number of target signals they were able to identify. The numbers reported by the participants were consistent with what is expected for this task, which indicates that the subjects understood the instructions. The whole task and recording session lasted 12 min on average.

3.2. Feature extraction

Hjorth features as biomarkers were extracted from the EEG signals of each of the three groups: control group (CTRL), PD patients who did not take their medicine (OFF), and PD patients who took their medicine (ON).

This mathematical parameter was firstly defined in Hjorth (1970), that proposed three measures called activity, mobility, and complexity in an attempt to establish parameters to describe EEG signals in terms of their amplitude and time range, but still related to the spectral shape of the signals. The activity feature is defined by the equation:

$$A_x = \sum_{n=1}^{N} \frac{[x[n] - \mu_x]^2}{N} \tag{1}$$

where $x[n]$ represents the discrete-time signal samples, μ_x its statistical mean, and N the signal length. Observing Eq. (1), one can see that activity corresponds to the estimated variance of a time series, which in turn represents the signal power.

The mobility feature is defined as the square root of the ratio between the variance of the signal first derivative and the variance of the signal, as shown by the equation:

$$M_x = \sqrt{\frac{\text{var}(\dot{x}[n])}{\text{var}(x[n])}} \tag{2}$$

where $\dot{x}[n]$ is the first derivative of the time signal and $\text{var}(\cdot)$ is the statistical variance operator. Analyzing Eq. (2), it can be inferred that the mobility depends exclusively on the shape of the curve represented by $x[n]$, and its value estimates the relative slope. In the spectral domain, the mobility corresponds to the standard deviation of the signal power spectrum.

Complexity is a dimensionless parameter, derived from the ratio between the mobility of the first derivative of the EEG signal and the mobility of the EEG signal, defined by:

$$C_x = \frac{M_{\dot{x}}}{M_x} \tag{3}$$

where M_x and $M_{\dot{x}}$ are the mobility of the signal and the mobility of its first derivate, respectively. The smallest possible value for complexity corresponds to one and is reached when analyzing a pure sinusoid. Therefore, complexity measures the degree of similarity of the EEG signal to a sine wave.

Although the Hjorth features have interpretations in the frequency domain, all three measures are calculated over the amplitudes of the time series, without the need for large computational cost, which makes them ideal for real-time analysis (Oh et al., 2014).

The Hjorth features were computed for all epochs of all channels for each participant EEG signal. For each channel, the normalized features were averaged over their epochs. All values were normalized in the range $(0, 1)$ using the equation:

$$\hat{x[n]} = \frac{x[n] - \min(x[n])}{\max(x[n]) - \min(x[n])} \tag{4}$$

in which x is the original feature and $\hat{x}$ its normalized value.

With this methodology, 180 Hjorth features were obtained for each subject, corresponding to the average activity, mobility, and complexity computed in each of the 60 channels. After that, Pearson correlation was performed between the extracted characteristics and a binary target vector representing the labels of two classes. The correlation was useful to describe how good the features were to distinguish these labels. Only the parameters that obtained a significance correlation $p \leq 0.05$ were selected for the classification step.

The behavior of the Hjorth features along the scalp was visualized by plotting the average values of the normalized activity, mobility, and complexity into topographical graphs. The channels that reached a p value lower than 0.05 were also plotted over the scalp, in order to show which brain areas were the most important to distinguish the groups.

3.3. Classification

In the classification step, the following binary categorizations were considered: CTRL × OFF, CTRL × ON, and OFF × ON. The classification was performed on MATLAB via the Classification Learner App.

The data selected in the previous step was split using a 5-fold cross-validation method, which means that data was randomly separated into five groups. During the training process, four of these groups were used to train the model and the one left was used to test the model. The process was repeated five times, thus in each repetition a group was left out of the training to be used as the test group. At the end of the five iterations, the obtained results were used to validate the model.

The classification algorithms used in this work were the linear SVM, the medium KNN, and the Random Forest. An SVM is a non-probabilistic, supervised binary classifier that analyzes a set of labeled data and, from that analysis, makes predictions on unlabeled data (Noble, 2006). In SVM, a kernel function maps the original set of points into a higher-dimensional vector space. These mapped points are then divided into two groups by an optimal separating hyper-plane. Based on this separating hyper-plane, the SVM assigns the unlabeled data to one of these two groups of points.

The KNN algorithm is one of the simplest machine learning models. The classifier determines which class the test sample belongs to by measuring the Euclidean distance to its nearest neighbor. If this neighbor belongs to the positive class, the classifier states that the test sample also belongs to the positive class. If the nearest neighbor is from the negative class, the classifier predicts that the test sample is also negative (Peterson, 2009). It is worth mentioning that this approach refers to the simplest KNN model, known as the 1-nearest neighbor rule.

Random Forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman, 2001). In this algorithm, the classification result is determined by the class selected by the most trees. This model has the advantage of showing more robustness against noise.

3.4. Model evaluation

The main tool for analyzing the performance of machine learning classifiers are measures extracted from the so-called confusion matrix (Witten & Frank, 2002). The primary diagonal of the confusion matrix accounts for the correct predictions of the classifier, where TP and TN indicate the true positives and the true negatives. The secondary diagonal shows the incorrect predictions of the classifier, with FP and FN being the false positives and the false negatives.

From the confusion matrix, other important metrics can be extracted, such as the true positive rate (TPR), the true negative rate (TNR), and the accuracy (ACC). These measures are obtained using Eq. (5), (6), and (7). TPR indicates the classifier's ability to predict the positive class, while TNR represents its ability to predict the negative class. Accuracy indicates the ratio between the amount of total correct predictions and the total number of predictions. Because it takes both positive and negative classes into account, accuracy is the main measure for evaluating classifiers.

$$TPR = \frac{TP}{TP + FN} \tag{5}$$

$$TNR = \frac{TN}{TN + FP} \tag{6}$$

$$ACC = \frac{TP + TN}{TP + FP + FN + TP} \tag{7}$$

The receiver operating characteristic (ROC) curve is another tool to assess the performance of a given computational classifier. The ROC graph evaluates the tradeoff between true positives and false positives by plotting the TPR against the false-positive rate (FPR, which corresponds to 1 - TNR) when the classifier threshold settings are changed (Fawcett, 2004).

Good performing classifiers have a ROC curve located in the upper half-plane defined by a line with a slope of 1, and the perfect classification (TPR and TNR equal to 1) is in the upper left corner. The area under the ROC curve (AUC) determines the ability of a classifier to discriminate between the two tested classes. An AUC equals to 0.5 refers to a classifier with no discriminating capability, while an AUC equals to 1.0 represents a perfect classifier (Hoo et al., 2017).

4. Results and discussion

4.1. Results

Average activity, mobility and complexity Hjorth features computed for different electrodes are shown in Figs. 2–4. The figures present how Hjorth features behave in a topographical, spatial way.

The extracted Hjorth features spotted major differences between control subjects and PD patients, while minor differences were found between the ON and OFF PD groups. A paired t-test was performed to investigate differences between groups. The average activity of the CTRL group was lower than the OFF ($p_t = 2.81 \times 10^{-15}$) and ON ($p_t = 2.80 \times 10^{-19}$) groups, especially in the electrodes located in the frontal, central, parietal, and occipital zones, as shown in Fig. 2. No significant differences were found on the average activity between ON and OFF groups ($p_t = 0.97$).

The values computed for mobility and complexity were also lower on the CTRL group (for mobility CTRL × OFF and CTRL × ON, respectively: $p_t = 1.22 \times 10^{-9}$, $p_t = 5.15 \times 10^{-6}$; for complexity CTRL × OFF and CTRL × ON, respectively: $p_t = 3.44 \times 10^{-17}$, $p_t = 2.27 \times 10^{-18}$), and indistinguishable between ON and OFF groups (for mobility: $p_t = 0.11$; for complexity: $p_t = 0.43$). For the mobility features, the main differences were found in the frontal and parietal zones, while the differences in parietal and occipital zones were more significant in complexity features, as presented in Figs. 3 and 4.

Fig. 5 indicates as red dots the channels that reached a p value lower than 0.05. The activity feature in frontal, central, and occipital zones was the most significant to distinguish CTRL × OFF and CTRL × ON. When used to differentiate OFF × ON, activity showed poor performance, with no significant correlation.

The mobility features extracted from some electrodes located in the central, parietal and occipital zones also reached p values lower than 0.05 for the CTRL × OFF and CTRL × ON situations, while for the OFF × ON case only one channel met the criteria, as observed in Fig. 6. As the previously analyzed parameters, the complexity features were significantly correlated with CTRL × OFF and CTRL × ON, especially in the central, parietal and frontal zones, and were not capable of distinguishing OFF × ON with good significance (Fig. 7).

Table 1 summarizes the electrodes with $p \leq 0.05$ for each situation. As said before, Hjorth features extracted from the frontal, central, parietal and occipital areas are good measures to discriminate CTRL × OFF and CTRL × ON. Meanwhile, only two features met the criteria of $p \leq 0.05$ to distinguish OFF × ON.

The selected Hjorth features, used as inputs for the machine learning algorithms, reached good results when classifying health subjects against PD patients. Figs. 8–10 compares the performance of all classifiers using the confusion matrices and the ROC curves computed for each situation. For the ROC curves, CTRL group was considered the positive class for the CTRL × OFF and CTRL × ON situations, and the OFF group was the positive class for the OFF × ON situation. The crosses in the graphs represent the optimal classifier point. Analyzing only

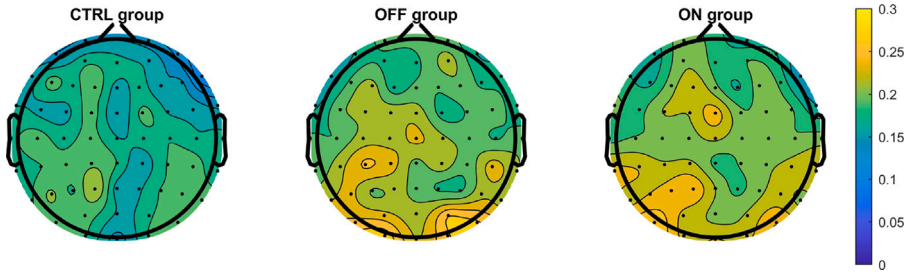

Fig. 2. Average activity computed for each of the three groups. The scale was adjusted in order to improve the visualization of the differences between groups. The mean activity on the CTRL group is lower when compared with the OFF and ON groups, especially in frontal, central, parietal, and occipital zones. No significant differences were found when comparing OFF and ON groups.

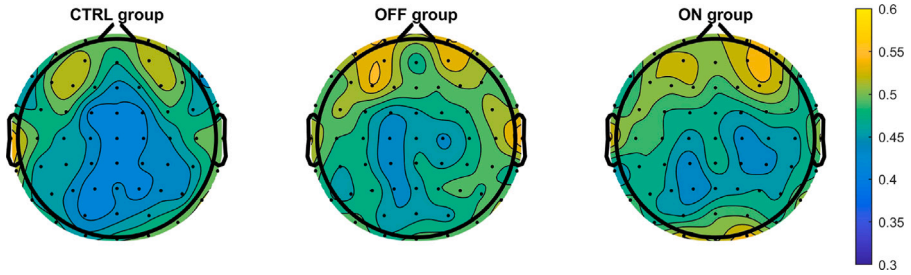

Fig. 3. Average mobility computed for each of the three groups. The scale was adjusted in order to improve the visualization of the differences between groups. The mean mobility on the CTRL group is lower when compared with the OFF and ON groups, especially in frontal and parietal zones. No significant differences were found when comparing OFF and ON groups.

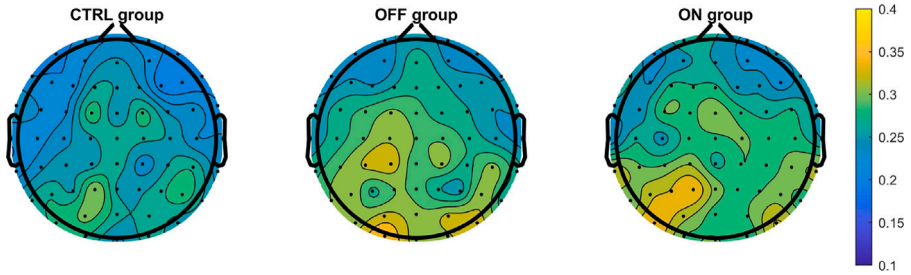

Fig. 4. Average complexity computed for each of the three groups. The scale was adjusted in order to improve the visualization of the differences between groups. The mean complexity on the CTRL group is lower when compared with the OFF and ON groups, especially in parietal, and occipital zones. No significant differences were found when comparing OFF and ON groups.

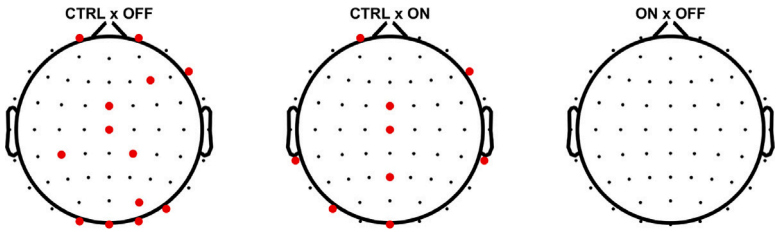

Fig. 5. Channels that reached $p \leq 0.05$ for the activity feature in each scenario. When distinguishing CTRL × OFF and CTRL × ON the electrodes located in frontal, central, and occipital zones were the most significantly correlated. No electrode met the criteria for the OFF × ON case.

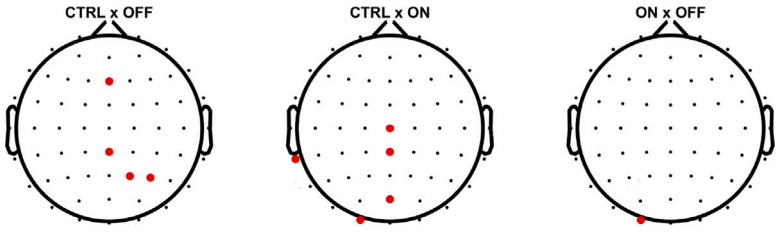

Fig. 6. Channels that reached $p \leq 0.05$ for the mobility feature in each scenario. When distinguishing CTRL × OFF and CTRL × ON the electrodes located in central, parietal, and occipital zones were the most significantly correlated. Only the electrode located in the occipital zone met the criteria for the OFF × ON case.

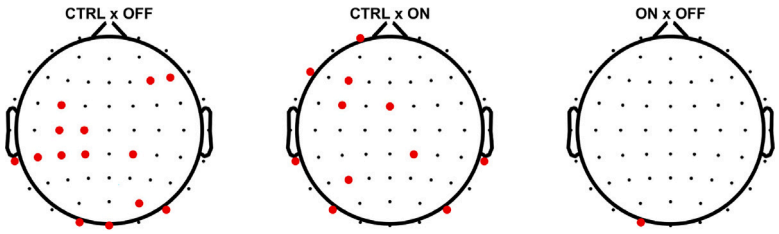

Fig. 7. Channels that reached $p \leq 0.05$ for the complexity feature in each scenario. When distinguishing CTRL × OFF and CTRL × ON the electrodes located in central, parietal and frontal zones were the most significantly correlated. Only the electrode located in the occipital zone met the criteria for the OFF × ON case.

Table 1
Electrodes where a significant difference in features was identified.

	Activity	Mobility	Complexity
CTRL × OFF	FP1, FP2, FCz, F4, F8, CP2, Cz, CP3, PO4, PO8, O1, O2, Oz.	Fz, CPz, P4, P2.	F4, F6, FC3, C1, C3, CP1, CP2, CP3, CP5, PO4, PO8, TP7, O1, Oz.
CTRL × ON	F8, FP1, FCz, Cz, Pz, PO7, TP7, TP8, Oz.	Cz, CPz, TP7, O1, POz.	F3, F7, FC3, FCz, FP1, CP2, P3, TP7, TP8, PO7, PO8.
OFF × ON		O1	O1

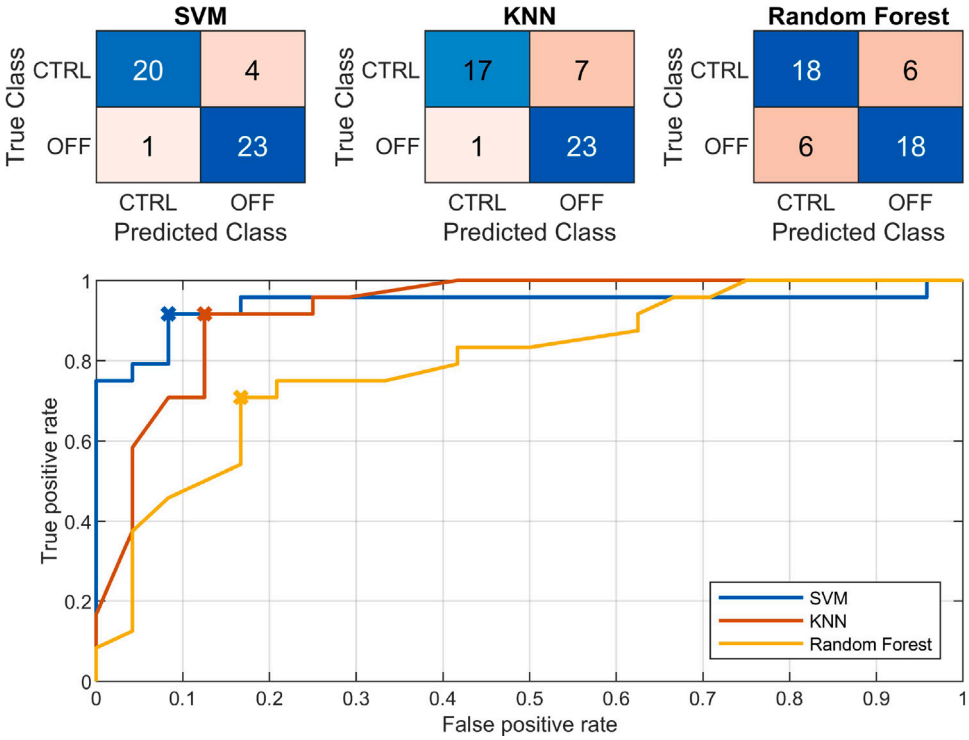

Fig. 8. Confusion matrices and ROC curve for each classifier when performing CTRL × OFF.

these curves and the confusion matrices, SVM and KNN outperformed Random Forest classifier in all three situations.

Table 2 summarizes the model evaluation metrics computed for each case. For the CTRL × OFF case, SVM reached an accuracy of 89.56%, KNN 83.56%, and Random Forest 76.22%. Applying a 5-by-2 paired F cross-validation test to verify if these accuracies are statistically different, the following p_F values were found: for SVM and KNN, $p_F = 0.7297$; for SVM and Random Forest, $p_F = 0.0333$; and for KNN and RF, $p_F = 0.6244$. Considering a significance level of 5%, only the accuracy of SVM and Random Forest are statistically different, which means that even though the classifiers were able to distinguish CTRL × OFF, it is not possible to state which one is the best, but only that SVM outperforms Random Forest.

When the algorithms were used to classify CTRL × ON, the KNN model achieved an ACC = 87.78%, the SVM model an ACC = 83,33%, and the Random Forest model an ACC = 68,67%. The p_F values obtained were: for SVM and KNN, $p_F = 0.4045$; for SVM and Random Forest, $p_F = 0.0265$; and for KNN and Random Forest, $p_F = 0.4712$. Again, this only shows that SVM outperforms RF. When classifying OFF × ON, SVM presented an accuracy of 66.22% and KNN 68.67%, while the RF made random predictions, resulting in an accuracy of 50.44%. It was also not possible to determine which classifier was the best, only that SVM was better than RF since the following p_F values were computed: for SVM and KNN, $p_F = 0.1957$; for SVM and Random Forest, $p_F = 0.0101$; and for KNN and RF, $p_F = 0.1304$.

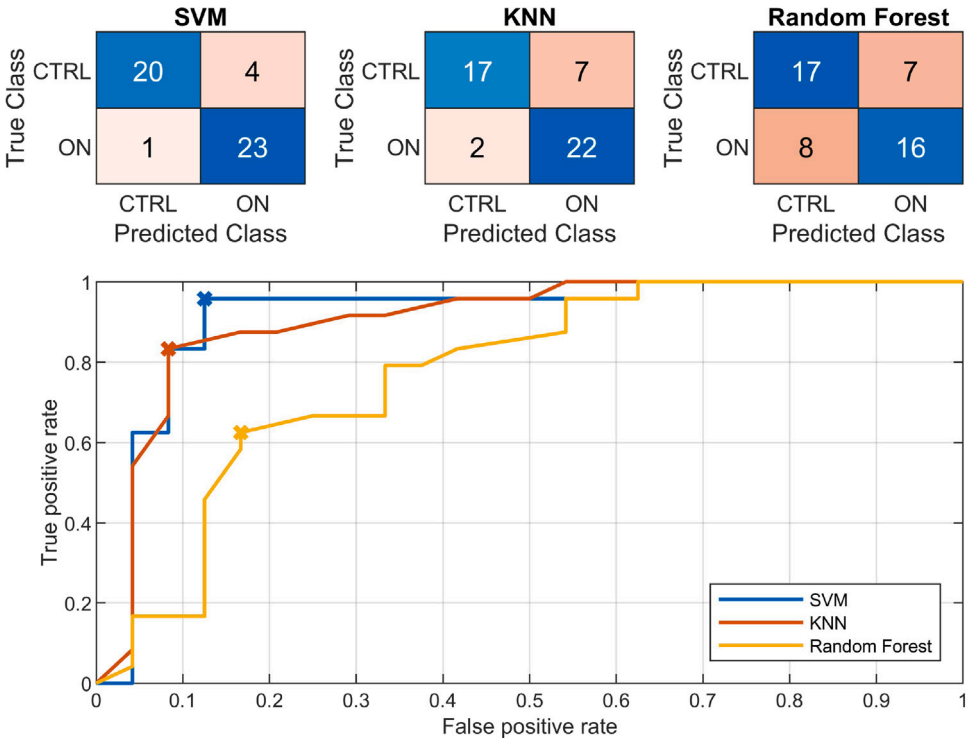

Fig. 9. Confusion matrices and ROC curve for each classifier when performing CTRL × ON.

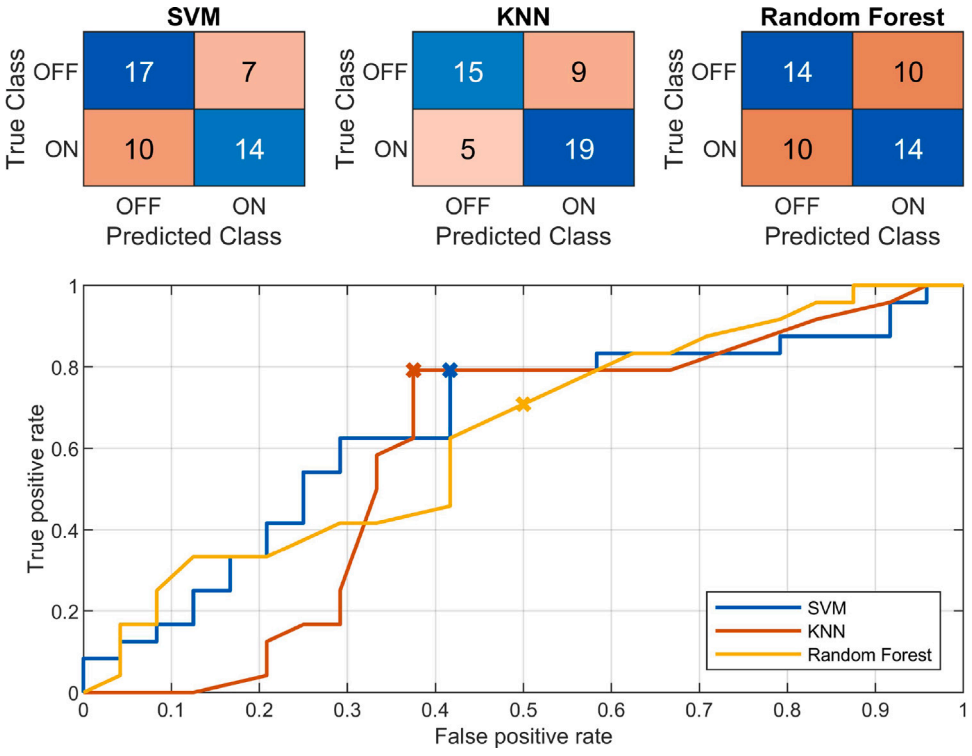

Fig. 10. Confusion matrices and ROC curve for each classifier when performing OFF × ON.

The high p_F values obtained with the 5-by-2 paired F cross-validation test indicate that the small sample size is a limitation and it probably explains why it was impossible to determine the best algorithm to classify the data. Nevertheless, the study shows that the classifiers are able to discriminate the CTRL × OFF and CTRL × ON situations with high accuracy.

4.2. Discussion

This work had the objective of investigating Hjorth features as biomarkers for Parkinson's disease. The paper showed that the activity and the complexity parameters outperformed the mobility ones. It was also shown that the values computed in the parietal, frontal, central, and occipital lobes were the most significant to differentiate healthy

Table 2
Classifiers performance. RF refers to the Random Forest classifier.

		TPR	TNR	ACC	AUC
CTRL × OFF	SVM	**0.9524**	**0.8519**	**0.8956**	0.9444
	KNN	0.9444	0.7667	0.8356	**0.9661**
	RF	0.7600	0.7826	0.7622	0.8299
CTRL × ON	SVM	0.9000	0.7857	0.8333	0.8767
	KNN	**0.9091**	**0.8462**	**0.8778**	**0.9106**
	RF	0.6957	0.6800	0.6867	0.7352
OFF × ON	SVM	0.6429	**0.7000**	0.6622	**0.6458**
	KNN	**0.7368**	0.6552	**0.6867**	0.5868
	RF	0.500	0.500	0.5044	0.4852

subjects and PD patients. The use of machine learning algorithms demonstrated that Hjorth features are able to characterize PD given that both the SVM and KNN models identified PD patients with high accuracy.

Several works have associated Parkinson's disease with altered activity in different brain regions. In Kubera et al. (2019), it was demonstrated that PD patients have a cortical thickness decrease loss in the frontal, cingulate, tempo-parietal, and occipital zones compared to matched healthy patients. Esmaeilzadeh et al. (2018), used MRI to show that the superior parietal region is of critical importance to diagnose PD. Silva et al. (2020) used an EEG dataset from the PRED+CT in which the data was collected with the patients in a total relaxation state to propose a method to detect early stages of PD using the cross-correlation function and machine learning algorithms. The authors observed that the cross-correlation computed in the right temporal, frontal, parietal, and occipital zones, when used as inputs for a feed-forward artificial neural network (fANN), detected PD with high accuracies. The founds in the present work follow the recent literature since it was able to detect clear differences between PD patients and healthy subjects in the parietal, frontal, central, and occipital zones.

The use of Hjorth features in the study of Parkinson's disease is somewhat rare, with no recent works at classifying PD patients using these parameters, as far as the author's knowledge. Nevertheless, other techniques were used to perform this task. Cavanagh et al. (2018) also used the same PRED+CT dataset to identify PD patients using the event-related potential. In that study, an SVM classifier was used with the first 50 terms of the fast Fourier transform of the signal immediately following a stimulus as inputs to the algorithm. The best accuracy was achieved when classifying CTRL × OFF groups, with an ACC = 82%. For the task of discriminating between CTRL × ON, the best accuracy was ACC = 75%, and when distinguishing OFF × ON, the classifier showed poor performance.

The results obtained in Cavanagh et al. (2018) were consistent with those found in the present work. In our study, the best accuracies were obtained when trying to discriminate the control group from PD patients. When classifying CTRL × OFF, the SVM algorithm showed the best performance, with an ACC = 89.56%. For the task of differentiating the CTRL × ON groups, the KNN reached an ACC = 87.78%, and when classifying OFF × ON, it obtained an ACC = 68.67%. These results indicate that Hjorth features seems to be better than fast Fourier transform coefficients to identify PD patients in this scenario. Even though the results showed an improvement from the previous work, the limited amount of data poses a limitation to identify the best classifier, and more studies are needed in order to compare the performance of the fast Fourier transform coefficients with the Hjorth features.

5. Conclusion

This work presented Hjorth features extracted from EEG signals as biomarkers for PD. These parameters showed a significant difference between healthy and PD patients, especially when computed in the parietal, frontal, central, and occipital zones. It was also possible to classify these two groups of patients using supervised machine learning algorithms (SVM and KNN).

With the analysis of the results, and the similar outcomes from other works, one can conclude that the proposed approach was efficient, providing a simple and efficient method to identify PD patients using EEG signals.

To further consolidate the reproducibility of the results achieved in this work, the use of Hjorth features as biomarkers for PD must be considered in larger datasets, which will also allow for the identification of the classifier with the best performance.

CRediT authorship contribution statement

Bruno Fonseca Oliveira Coelho: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft. **Ana Beatriz Rodrigues Massaranduba:** Conceptualization, Methodology, Writing – review & editing. **Carolline Angela dos Santos Souza:** Conceptualization, Methodology. **Giovanni Guimarães Viana:** Conceptualization, Methodology. **Ivani Brys:** Conceptualization, Supervision, Writing – review & editing. **Rodrigo Pereira Ramos:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data is available online in the PRED+CT repository.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eswa.2022.118772.

References

Barcelon, E. A., Mukaino, T., Yokoyama, J., Uehara, T., Ogata, K., Kira, J.-i., & Tobimatsu, S. (2019). Grand total EEG score can differentiate Parkinson's disease from Parkinson-related disorders. *Frontiers in Neurology*, *10*, 398.

Bhat, S., Acharya, U. R., Hagiwara, Y., Dadmehr, N., & Adeli, H. (2018). Parkinson's disease: Cause factors, measurable indicators, and early diagnosis. *Computers in Biology and Medicine*, *102*, 234–241.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, 16.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Cavanagh, J. F., Kumar, P., Mueller, A. A., Richardson, S. P., & Mueen, A. (2018). Diminished EEG habituation to novel events effectively classifies Parkinson's patients. *Clinical Neurophysiology*, *129*(2), 409–418.

Cavanagh, J. F., Napolitano, A., Wu, C., & Mueen, A. (2017). The patient repository for EEG data + computational tools (PRED+CT). *Frontiers in Neuroinformatics*, *11*, 67.

Chen, X., Wang, Z. J., & McKeown, M. (2016). Joint blind source separation for neurophysiological data analysis: Multiset and multimodal methods. *IEEE Signal Processing Magazine*, *33*(3), 86–107.

de Oliveira, A. P. S., de Santana, M. A., Andrade, M. K. S., Gomes, J. C., Rodrigues, M. C., & dos Santos, W. P. (2020). Early diagnosis of Parkinson's disease using EEG, machine learning and partial directed coherence. *Research on Biomedical Engineering*, *36*(3), 311–331.

Esmaeilzadeh, S., Yang, Y., & Adeli, E. (2018). End-to-end parkinson disease diagnosis using brain mr-images by 3d-cnn. arXiv preprint arXiv:1806.05233.

Fawcett, T. (2004). ROC graphs: Notes and practical considerations for researchers. *Machine Learning*, *31*, 1–38.

Fichtenholtz, H. M., Dean, H. L., Dillon, D. G., Yamasaki, H., McCarthy, G., & LaBar, K. S. (2004). Emotion–attention network interactions during a visual oddball task. *Cognitive Brain Research*, *20*(1), 67–80.

Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, *29*(3), 306–310.

Hoo, Z. H., Candlish, J., & Teare, D. (2017). What is an ROC curve? *Emergency Medicine Journal*, *34*(6), 357–359. http://dx.doi.org/10.1136/emermed-2017-206735, arXiv: https://emj.bmj.com/content/34/6/357.full.pdf.

Jankovic, J. (2008). Parkinson's disease: Clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*(4), 368–376.

Jenke, R., Peer, A., & Buss, M. (2014). Feature extraction and selection for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, *5*(3), 327–339.

Klassen, B., Hentz, J., Shill, H., Driver-Dunckley, E., Evidente, V., Sabbagh, M., Adler, C., & Caviness, J. (2011). Quantitative EEG as a predictive biomarker for Parkinson disease dementia. *Neurology*, *77*(2), 118–124.

Koch, M., Geraedts, V., Wang, H., Tannemaat, M., & Bäck, T. (2019). Automated machine learning for EEG-based classification of Parkinson's disease patients. In *2019 IEEE international conference on big data* (pp. 4845–4852). IEEE.

Kubera, K. M., Schmitgen, M. M., Nagel, S., Hess, K., Herweh, C., Hirjak, D., Sambataro, F., & Wolf, R. C. (2019). A search for cortical correlates of trait impulsivity in Parkinson s disease. *Behavioural Brain Research*, *369*, Article 111911.

Moisello, C., Blanco, D., Lin, J., Panday, P., Kelly, S. P., Quartarone, A., Di Rocco, A., Cirelli, C., Tononi, G., & Ghilardi, M. F. (2015). Practice changes beta power at rest and its modulation during movement in healthy subjects but not in patients with Parkinson's disease. *Brain and Behavior*, *5*(10), Article e00374.

Noble, W. S. (2006). What is a support vector machine? *Nature biotechnology*, *24*(12), 1565–1567.

Oh, S. L., Hagiwara, Y., Raghavendra, U., Yuvaraj, R., Arunkumar, N., Murugappan, M., & Acharya, U. R. (2018). A deep learning approach for Parkinson's disease diagnosis from EEG signals. *Neural Computing and Applications*, 1–7.

Oh, S.-H., Lee, Y.-R., & Kim, H.-N. (2014). A novel EEG feature extraction method using Hjorth parameter. *International Journal of Electronics and Electrical Engineering*, *2*(2), 106–110.

Patil, A., Deshmukh, C., & Panat, A. (2016). Feature extraction of EEG for emotion recognition using Hjorth features and higher order crossings. In *2016 Conference on advances in signal processing* (pp. 429–434). IEEE.

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, *4*(2), 1883.

Petrantonakis, P. C., & Hadjileontiadis, L. J. (2009). Emotion recognition from EEG using higher order crossings. *IEEE Transactions on Information Technology in Biomedicine*, *14*(2), 186–197.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(10), 1175–1191.

Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., Schrag, A.-E., & Lang, A. E. (2017). Parkinson disease. *Nature Reviews Disease Primers*, *3*(1), 1–21.

Potter, D. D., Bassett, M. R., Jory, S. H., & Barrett, K. (2001). Changes in event-related potentials in a three-stimulus auditory oddball task after mild head injury. *Neuropsychologia*, *39*(13), 1464–1472.

Sakar, B. E., Isenkul, M. E., Sakar, C. O., Sertbas, A., Gurgen, F., Delil, S., Apaydin, H., & Kursun, O. (2013). Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*, *17*(4), 828–834.

Shetty, S., & Rao, Y. (2016). SVM based machine learning approach to identify Parkinson's disease using gait analysis. In *2016 International conference on inventive computation technologies, Vol. 2* (pp. 1–5). IEEE.

Silva, G., Alves, M., Cunha, R., Bispo, B. C., & Rodrigues, P. M. (2020). Parkinson disease early detection using EEG channels cross-correlation. *International Journal of Applied Engineering Research*, *15*(3), 197–203.

Vanegas, M. I., Blangero, A., Galvin, J. E., Di Rocco, A., Quartarone, A., Ghilardi, M. F., & Kelly, S. P. (2019). Altered dynamics of visual contextual interactions in Parkinson's disease. *Npj Parkinson's Disease*, *5*(1), 1–8.

Vanegas, M. I., Ghilardi, M. F., Kelly, S. P., & Blangero, A. (2018). Machine learning for EEG-based biomarkers in Parkinson's disease. In *2018 IEEE international conference on bioinformatics and biomedicine* (pp. 2661–2665). IEEE.

Wahid, F., Begg, R. K., Hass, C. J., Halgamuge, S., & Ackland, D. C. (2015). Classification of Parkinson's disease gait using spatial-temporal gait features. *IEEE Journal of Biomedical and Health Informatics*, *19*(6), 1794–1802.

Wang, T., Shoaran, M., & Emami, A. (2018). Towards adaptive deep brain stimulation in Parkinson's disease: Lfp-based feature analysis and classification. In *2018 IEEE international conference on acoustics, speech and signal processing* (pp. 2536–2540). IEEE.

Witten, I. H., & Frank, E. (2002). Data mining: Practical machine learning tools and techniques with java implementations. *Acm Sigmod Record*, *31*(1), 76–77.

Yao, L., Brown, P., & Shoaran, M. (2018). Resting tremor detection in Parkinson's disease with machine learning and Kalman filtering. In *2018 IEEE biomedical circuits and systems conference* (pp. 1–4). IEEE.

Yuvaraj, R., Acharya, U. R., & Hagiwara, Y. (2018). A novel Parkinson's disease diagnosis index using higher-order spectra features in EEG signals. *Neural Computing and Applications*, *30*(4), 1225–1235.