

EEG-based PD Classification Model Coupled with Machine Learning

Zhen Bian

Department of Microelectronics Science and Engineering, Sun Yat-sen University,
Guangdong, China
bianzh5@mail2.sysu.edu.cn

Abstract. One of the most prevalent neurological brain diseases is Parkinson's disease, which can be diagnosed a long time ago with a variety of clinical methods. In recent years, it has been common practice to use Electroencephalography (EEG) signal analysis to identify dementia in its early stages because of its high speed, low cost, and accessibility. Many novel methods which apply EEG to the diagnosis of Parkinson's disease are shown to be simple and effective. Recent years have seen the development of EEG signal processing as a key technique for researchers to gather appropriate features for Parkinson's disease diagnosis. In this study, a novel system was created for computer-aided diagnosis that is capable of extracting features from EEG signals and discriminating patients affected by Parkinson's disease. After pre-processing the EEG data, the Butterworth filter has been used to decompose the signals into four frequency sub-bands. Welch's PSD features are then extracted as the input of supervised machine learning methods-the k-Nearest Neighbor (KNN) to classify EEG features into Parkinson's disease (PD) and healthy controls (HC). The 10-fold cross-validation has been employed to validate the performance of this model. The results achieve 98.82% accuracy, 99.19% sensitivity, and 91.77% specificity, respectively. The acquired findings demonstrate the validity of our strategy and that our diagnosis method is improved when compared to earlier research. At last, this novel method may be a supplementary tool for the clinical diagnosis of Parkinson's disease.

Keywords: PD detection; EEG; PSD; Classification; Machine Learning

1. Introduction

One of the age-related and neurodegenerative disorders that is prevalent in many nations is Parkinson's disease (PD). It is brought on by a decrease in the basal ganglia portion of the brain's ability to secrete dopamine. Aging is the main factor causing dopamine secretion to decrease in age-related brain neurons. The brain state is significantly impacted by the neurons' decreased ability to create dopamine. As a result, PD affects elderly adults more frequently. According to studies, this disease affects about 0.3% of the world's population. 80 percent or more of PD patients are 60 years of age or older. In addition, a different study found that men have a 1.5 times higher risk of having PD. Furthermore, this condition is most common in adults aged 85 to 89. Patients with PD may experience emotional, cognitive, and behavioral difficulties. Slower movements, trouble speaking, and shaking hands and feet while at rest are the most typical symptoms among PD patients. The daily lives of PD sufferers are negatively impacted by this circumstance. Parkinson's disease diagnosis has always been based on

clinical and anamnesis data. Slower motions and hand tremors are two of PD's most significant symptoms. A preliminary diagnosis of PD can also be made using mood disorders including sadness and sleep disorders involving rapid eye movements (REM). The diagnosis of PD is based on a variety of physiological and pathological data, however, these results do not completely separate healthy from diseased people [1].

Recently, a variety of machine-learning and signal-processing techniques for EEG feature extraction as well as classification have been presented and investigated with the goal of creating an early diagnosis system that can evaluate brain signals automatically and aid neurologists in the early detection of neurological illnesses like Parkinson's disease, Alzheimer disorders, Autism Spectrum Disorders (ASD), and Epilepsy Disorders (ED). As a result, many scientists are already developing computer models that can detect Alzheimer's disease by analyzing patient brain signals.

Traditional PD detection methods are upsetting and time-costly. Electroencephalography (EEG) is a quick, affordable, and non-invasive method of collecting brain signals; nevertheless, its interpretation necessitates the visual examination, which can be time costly and dependent on the performer's level of skill. Furthermore, once the EEG signal is lengthy, manual inspection takes a long time and is fraught with inaccuracies due to the existence of artifacts mixed with the signal. The automated model based on EEG signal processing coupled with supervised machine learning techniques has emerged as a significant research area to assist physicians with the challenging process of PD detection [2].

The international 10-20 system is the most widely used configuration for the electrodes that are placed on the scalp as part of the EEG data sampling technique. For the purpose of recording the electrical activity of the cerebral cortex, electrodes capture the postsynaptic biopotentials of all neurons with the same spatial orientation. Biopotentials are the building blocks of the raw signals and are sampled either in monopolar mode with a reference electrode or in bipolar mode by various electrode pairs. Pre-processing procedures then use the band-pass filter to remove noise and eliminate artifacts from the raw data. All pre-processing procedures, in essence, convert raw signals to EEG signals [3].

EEG signal characteristics from various domains such as frequency and time are extrapolated by all feature extraction techniques. Following that, models based on nonlinear or linear systems can be developed and validated using statistical or machine learning methods to identify PD from HC or normal aging.

This paper describes a method that can help early predict PD from HC by utilizing feature extraction methods coupled with supervised machine learning techniques. Our idea is to use the extracted features for classification. For this purpose, Welch's method was applied to extract features, then make the extracted Power Spectral Density (PSD) features the input of the kNN classifier.

2. Method

This section has first provided a description of the EEG dataset that was used. Afterward, the suggested technique for handling the EEG data has also been explained. Pre-processing, feature extraction, and classification are the three primary procedures that are taken in order to construct a system for diagnosing Parkinson's disease, as illustrated in figure 1. Firstly, the collected EEG signals were imported and get processed using pre-processing methods to eliminate the impact of noise from artifacts and make the signals re-referenced. Meanwhile, the TP9 and TP10 channels were used to re-reference, and there were 61 available channels after re-referencing. The frequency properties of the EEG signals were typically represented by four frequency sub-bands, that is Delta(0-4Hz), Theta(4-8Hz), Alpha(8-13Hz), and Beta(13-30Hz) [4]. Consequently, an FIR filter was used to limit the frequency of the re-referenced signals between 0.1Hz and 30Hz to contain all frequencies from the

four sub-bands and get rid of other useless signals. The filtered signal was then decomposed into frequency sub-bands using the 3rd-order Butterworth filter. Following that, the feature vectors were obtained from 61 channels by calculating the PSD of each EEG frequency sub-band. The dimension of the extracted feature is 763×244 (epochs \times features). Finally, the extracted features were imported into the k-Nearest Neighbor (kNN) classifier to classify the EEG data. The performance of this system was assessed using the 10-fold cross-validation. MATLAB software tools have been used to verify the proposed method. Each stage has been given and covered in more detail in the subsections that follow.

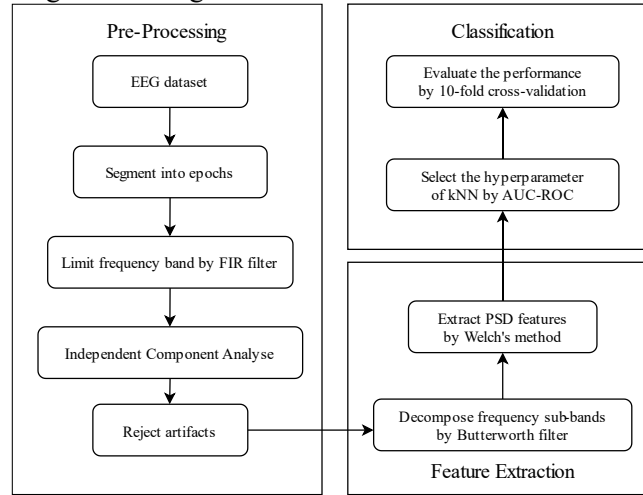


Figure 1. The three main steps in the proposed model.

2.1. Data Description

An open EEG dataset was utilized in this study to automatically identify PD. The information about this dataset is provided below. The EEG dataset is composed of 12 PD patients and 12 healthy controls from the University of Iowa (UI; Iowa City, Iowa). All the data were collected in resting condition with eyes open. With a 64-channel Brain Vision system, EEG was recorded from sintered Ag/AgCl electrodes spanning the frequency range of 0.1 to 100 Hz at a sampling rate of 500 Hz, with the online reference set to channel Pz as the baseline. Each recording of electrical brain activity lasts about 200s. The recording EEG signals were segmented into epochs of 6-s, in hopes of accomplishing optimal classification performance [5]. There are 763 EEG epochs which include 394 HC and 369 PD.

2.2. Pre-processing

The FIR filter was used in the pre-processing step to filter the EEG recordings at the pass band's 30Hz high edge and 0.1Hz lower edge. Then the datasets were re-referenced using TP9 and TP10 channels as references.

After re-referencing, Independent Component Analysis (ICA) decomposition was used to remove the artifacts from the data. The main problem in analyzing the EEG signal is the mixture of the different kinds of interference artifacts and the EEG data during the recording stage. The noise from artifacts such as eye movements, blinks, muscles, and heart activities can severely contaminate EEG activity. ICA decomposition can separate multiple artifacts from EEG data by linear decomposition, then the artifacts are rejected to get a clean dataset.

2.3. Feature Extraction

There are two primary processes in the feature extraction process: (i) obtain the sub-bands of each epoch and (ii) extract PSD from sub-bands using Welch's method in the frequency domains.

2.3.1. Obtain the sub-bands using the Butterworth filter.

Delta (0–4 Hz), Theta (4–8 Hz), Alpha (8–13 Hz), and Beta (13–30 Hz) were used as the four frequency sub-bands to divide the data of brain activity. The 3rd-order Butterworth bandpass filter was used to obtain four frequency sub-bands by setting the different lower edge and higher edge. The Butterworth filter was chosen because it has a linear response compared with others [6]. The Butterworth filter responds in a manner that is as flat as possible, with a roll-off of minus 20 dB per pole and no passband ripple. The Butterworth filter's frequency response falls off into zero in the stopband and is optimally flat in the passband.

2.3.2. Extract PSD feature by Welch's method

The features of each data were extracted by finding Power Spectral Density using Welch's method. The weighted sum of the periodograms of overlapping data windows is used in Welch's method, a non-parametric computing technique [7].

Welch's PSD was computed as follows:

Step 1: The input vector was partitioned into K segments. Each segment had M points.

Step 2: Apply the windowed Discrete Fourier Transform (DFT) to each segment where k equals 1 to K, along with the window function $\omega[m]$.

$$X_k(f) = \sum_m x[m]\omega[m]\exp(-j2\pi fm) \quad [1]$$

Step 3: Compute the modified periodogram value $P_k(f)$ for each segment.

$$P_k(f) = \frac{1}{W} |X_k(f)|^2 \quad [2]$$

where

$$W = \sum_{m=0}^M \omega^2[m] \quad [3]$$

Step 4: Obtain the Welch's PSD by computing the average of the periodogram value.

$$S(f) = \frac{1}{K} \sum_{k=1}^K P_k(f) \quad [4]$$

The input vector is often segmented into the longest segments to get as near to 8 segments with 50% overlap as possible, but not more. Hamming windows are used to window each section. The coefficients for the sub-bands of each channel obtained after computing Welch's PSD were used as features that served as the input of the classifier. As a result, each epoch had a 244-length feature vector.

2.4. Design of Classifier

The k-Nearest Neighbor (kNN) classifier was used to classify the data along with the extracted PSD features since it has shown fast classification speed and high accuracy [8]. The kNN classifier is constructed based on the distance approach between different inputs. The main architecture of the kNN classifier is shown in figure 2. The number of nearest neighbors that are considered while determining a new instance's class is defined by the hyperparameter 'k'. The noise could have a large impact on the kNN classifier when the value of the hyperparameter is small because the instance will be more sensitive to its neighbors, which means if its neighbors are noisy points the classification could be less accurate. While the large value of the hyperparameter means the computation of classification is expensive. In this paper, the value of the hyperparameter is carefully selected to obtain the balance between classification accuracy and computation time by calculating the Area Under the Receiver Operating Characteristic curve (AUR-ROC) for kNN with different hyperparameters.

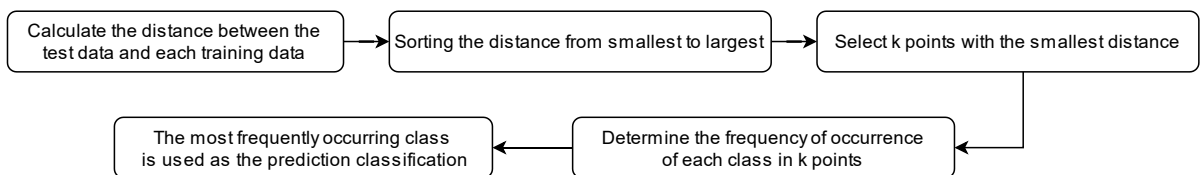


Figure 2. The main architecture of the kNN classifier.

2.5. Performance Evaluation and Cross-Validation

The k-fold cross-validation algorithm has been used in this paper to obtain the hyperparameter of the kNN classifier. All EEG features were equally and randomly segmented into k subsets. The remaining subsets were utilized for training, and one subset was chosen for the testing (validation) step. To make each subset used once for the testing process, this procedure will be repeated k times (k-fold). In this paper, k was chosen to be ten. The method which evaluates the performance of the system by plotting the Receiver Operating Characteristic curve (ROC) was used in this paper. The accuracy (ACC), true negative rate (TNR), and true positive rate (TPR) were calculated as well to evaluate the performance of classification.

3. Result

The environment of MATLAB (R2021b) was used to develop the proposed EEG signal classification system. A laptop with 16 GB of RAM and a 5 GHz Intel i7-10750H processor was used.

To find the optimal kNN hyperparameter, 10-fold cross-validation was used on this system first. As a result, the hyperparameter values between 1 and 15 were chosen for the 10-fold cross-validation. The accuracy (ACC), sensitivity (TPR), specificity (TNR), and AUC-ROC were computed for each value of the hyperparameter to evaluate its performance, as shown in table 1.

Table 1. The performance of kNN with different hyperparameters.

Hyperparameter	Accuracy	Sensitivity	Specificity	AUR-ROC
1	98.98±0.1	98.61±0.1	99.38±0.1	98.98±0.1
2	97.89±0.1	96.42±0.1	99.47±0.1	99.69±0.1
3	98.88±0.1	98.88±0.1	98.9±0.1	99.67±0.1
4	98.44±0.1	97.59±0.1	99.35±0.1	99.79±0.1
5	98.89±0.1	98.98±0.1	98.8±0.1	99.81±0.1
6	98.55±0.1	98.27±0.1	98.87±0.1	99.91±0.09
7	98.82±0.1	99.19±0.1	98.44±0.1	99.92±0.08
8	98.74±0.1	98.85±0.1	98.64±0.1	99.92±0.08
9	98.45±0.1	98.99±0.1	97.88±0.1	99.88±0.1
10	98.35±0.1	98.58±0.1	98.11±0.1	99.88±0.1
11	98.11±0.1	98.76±0.1	97.43±0.1	99.85±0.1
12	98.25±0.1	98.65±0.1	97.85±0.1	99.83±0.1
13	97.92±0.1	98.81±0.1	96.99±0.1	99.82±0.1
14	97.99±0.1	98.68±0.1	97.28±0.1	99.81±0.1
15	97.64±0.1	98.81±0.1	96.42±0.1	99.79±0.1

The Area Under the Receiver Operating Characteristic curve (AUR-ROC) reached the max value when the hyperparameter was selected as seven. The highest AUR-ROC value means the best classification ability. Therefore, the hyperparameter of the kNN classifier is determined as seven. The confusion matrix of the kNN classifier is shown in table 2. And the Receiver Operating Characteristic (ROC) curve was also plotted as shown in Figure 3 when the hyperparameter equals seven.

Table 2. The confusion matrix of the kNN classifier when the hyperparameter equals seven.

Class	HC	PD
HC	51.25%	0.79%
PD	0.39%	47.57%

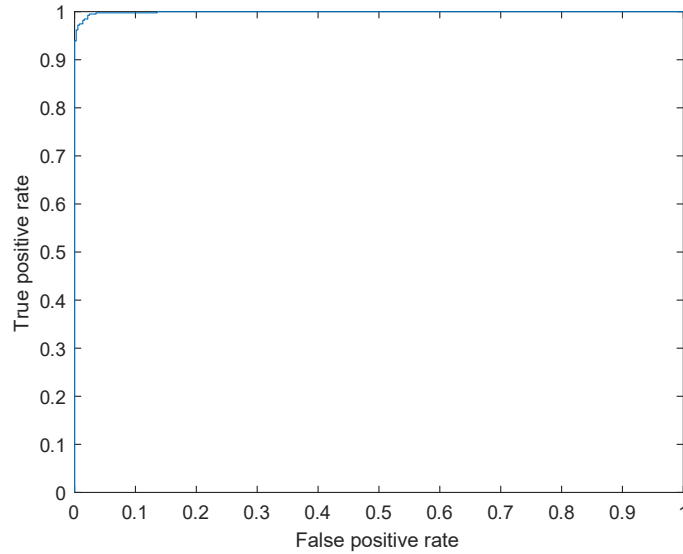


Figure 3. The ROC curve of the kNN when the hyperparameter equals seven.

4. Discussion

This paper presents a machine learning-based EEG signal classification system. This method quickly decomposes the filtered EEG signals between 0.1Hz and 30 Hz into a few epochs to extract classification features. In our model, PSD features are extracted after the pre-processing stage. By deploying Welch's method, each sample was extracted with 244 features from the EEG signals. All features are selected as the input of the kNN classifier. The hyperparameter of kNN was carefully selected as seven by employing the 10-fold cross-validation to obtain the highest performance. The performance shows that these features are efficient for classification and this model is useful for PD detection.

To evaluate the effectiveness and superiority, this proposed classification system is compared with existing classification techniques. In 118 PD patients, Betrouni et al. utilized FFT to identify the various stages of the disease [9]. Employing SVM and kNN classifiers, their method obtained 84% and 88% accuracy, respectively. In a dataset consisting of 10 PD and 10 HC participants, Naghsh et al. exploited spectral characteristics acquired from the FFT approach to detect PD [10]. Employing k-means clustering algorithms and fine-Gaussian SVM, their model had an accuracy of 95%. For the automatic detection of PD, Oh et al. employed a thirteen-layered one-dimensional CNN; the accuracy, sensitivity, and specificity values of their method were 84.71%, 84.72%, and 91.77%, respectively [11]. The accuracy of our presented model was 98.82%, the sensitivity was 99.19%, and the specificity was 98.44%.

The main advantages of our method are as follows:

- The accuracy of the model is comparable to the existing classification methods, which has been validated by 10-fold cross-validation.
- The model is easy to accomplish as the number of features is small but efficient for classification.
- The model is mostly automatic since it only needs a few manual adjustments.
- The selection of the hyperparameter and the length of segmented epochs makes the classification fast, robust, and accurate.

The following are some drawbacks of our model:

- The used dataset only contains 20 subjects.
- The filtering method used in the pre-processing stage is inefficient.

In the future, the novel filter method which can eliminate most of the noise will be added in the pre-processing step. Bigger and more diverse datasets are planned to validate this model. And the model is planned to test the detection and classification of other neurological disorders, such as Epilepsy Disorders (ED), Parkinson's disease (PD), and Autism Spectrum Disorder (ASD), in order to accomplish a unified EEG-based neurological disorder detection model.

5. Conclusion

An accurate, automatic, and fast EEG-based PD classification model is proposed in this paper. The system first segments the EEG signals into epochs of 6-s length. The frequency of the signals is then restricted to between 0.1Hz and 30Hz using the FIR filter. Four sub-bands (Delta, Theta, Alpha, Beta, and Gamma) are extracted by the Butterworth filter. After the pre-processing stage, this model utilizes Welch's PSD features along with a kNN classifier in order to automatically classify PD.

The 10-fold cross-validation validation has been applied to this model aiming to show the superiority of its performance. Our model attained an accuracy of 98.82%, a sensitivity of 99.19%, and a specificity of 98.44% respectively. Moreover, the results of our model were compared to other existing classification models. The results clearly illustrate that our suggested EEG-based PD classification model may be utilized in real-world PD detection missions. Our proposed prototype model can be used in real-time PD detection and is completely portable.

In the future, the developed EEG-based classification model could be used to automate the identification of other neurological illnesses such as epilepsy, autism spectrum disorder, and Parkinson's disease.

References

- [1] Vanegas M I, Ghilardi M F, Kelly S P, et al. Machine learning for EEG-based biomarkers in Parkinson's disease[C]//2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2018: 2661-2665.
- [2] Perez-Valero E, Lopez-Gordo M A, Morillas C, et al. A review of automated techniques for assisting the early detection of Alzheimer's disease with a focus on EEG[J]. Journal of Alzheimer's disease, 2021, 80(4): 1363-1376.
- [3] Sharma N, Kolekar M H, Jha K, et al. EEG and cognitive biomarkers based mild cognitive impairment diagnosis[J]. Irbm, 2019, 40(2): 113-121.
- [4] Han C X, Wang J, Yi G S, et al. Investigation of EEG abnormalities in the early stage of Parkinson's disease[J]. Cognitive neurodynamics, 2013, 7: 351-359.
- [5] Yuvaraj R, Murugappan M, Ibrahim N M, et al. Optimal set of EEG features for emotional state classification and trajectory visualization in Parkinson's disease[J]. International Journal of Psychophysiology, 2014, 94(3): 482-495.
- [6] Daud S S, Sudirman R. Butterworth bandpass and stationary wavelet transform filter comparison for electroencephalography signal[C]//2015 6th international conference on intelligent systems, modelling and simulation. IEEE, 2015: 123-126.
- [7] Francis M J N, Keran M P, Chetan R, et al. EEG-Controlled Robot navigation using Hjorth parameters and Welch-PSD[J]. International Journal of Intelligent Engineering and Systems, 2021, 14(4): 231-240.

- [8] Barua P D, Dogan S, Tuncer T, et al. Novel automated PD detection system using aspirin pattern with EEG signals[J]. Computers in Biology and Medicine, 2021, 137: 104841.
- [9] Betrouni N, Delval A, Chaton L, et al. Electroencephalography - based machine learning for cognitive profiling in Parkinson's disease: Preliminary results[J]. Movement Disorders, 2019, 34(2): 210-217.
- [10] Naghsh E, Sabahi M F, Beheshti S. Spatial analysis of EEG signals for Parkinson's disease stage detection[J]. Signal, Image and Video Processing, 2020, 14(2): 397-405.
- [11] Oh S L, Hagiwara Y, Raghavendra U, et al. A deep learning approach for Parkinson's disease diagnosis from EEG signals[J]. Neural Computing and Applications, 2020, 32: 10927-10933.