



Semnan University

*Journal of Modeling & Simulation in Electrical & Electronics
Engineering (MSEEE)*

Journal homepage: <https://mseee.semnan.ac.ir/>

ISSN: 2821-0786



Presenting a Method for Detecting Parkinson's Disease from Electroencephalography (EEG) Signals Using Machine Learning Techniques

Seyed Ali Sadatnoori^{1*} and Hakim Golshahi²

Abstract--Parkinson's disease (PD) is an incurable neurological disorder caused by a decrease in dopamine levels in the human brain. This paper investigates and analyzes EEG signals from patients with Parkinson's disease and healthy individuals. To simulate the proposed method, the Timing Interval PD database available on the UNM portal was used, which includes EEG signals from 37 healthy volunteers and 92 patients with Parkinson's disease. The average age of the volunteers in the healthy group was 71.3 years, and in the patient group was 68.7 years. EEG signals were recorded with a sampling frequency of 500 Hz and 63 channels. To increase the processing quality, the signals have undergone preprocessing steps, including filtering, noise removal, normalization, and segmentation. In the filtering step, a band-pass filter with a frequency range of 1 to 30 Hz has been used to remove additional noise. Also, the signals were normalized using the z-score normalization method and divided into ten-second segments. Four classification methods, including Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree (DT), and Random Forest (RF), have been used. The proposed method has achieved the best performance of 96.80. The performance of the models has been evaluated using accuracy, sensitivity, and specificity indices.

Index Terms: Electroencephalogram signal, machine learning, Parkinson's disease, feature reduction, classification.

I. INTRODUCTION

Parkinson's disease (PD) is an incurable neurological disorder caused by a decrease in dopamine levels in the human brain. Dopamine is a neurotransmitter that helps send basic messages and is responsible for controlling movement and coordination. Dopamine levels decrease when ganglion cells die or become damaged. Unfortunately, many patients are not diagnosed in the early stages of the disease, which leads to a permanent, incurable neurological disorder. This makes PD untreatable and often fatal in later stages. According to the World Health Organization (WHO), approximately 10 million people worldwide have PD. Currently, tests for the disease are expensive and not highly accurate. This highlights the urgent need for a low-cost, efficient, and accurate early diagnostic method that allows for timely treatment before the disease becomes severe. However, doctors use a combination of symptoms and diagnostic tests to identify the disease. Studies show that PD can begin before motor symptoms appear, and about 90% of PD patients experience vocal disorders. Symptoms of Parkinson's disease may include tremors, restricted or slow movement (bradykinesia), impaired balance, impaired posture, involuntary movements (dyskinesia), muscle stiffness, and changes in

Received; 2026-05-31

Revised; 2026-07-25

Accepted; 2026-09-05

1. Department of Electrical Engineering, Sho. C, Islamic Azad University, Shoushtar, Iran.

2. Department of Physics, Sho. C, Islamic Azad University, Shoushtar, Iran.

*Corresponding author Email: sadatnoori7218@iau.ac.ir.

Cite this article as:

Sadatnoori, S. A., and Golshahi, H. (2026). Presenting a Method for Detecting Parkinson's Disease from Electroencephalography (EEG) Signals Using Machine Learning Techniques. *Journal of Modeling & Simulation in Electrical & Electronics Engineering (MSEEE)*. Semnan University Press. 7 (2), 1-9.
DOI: <https://doi.org/10.22075/mseee.2026.41427.1280>

© 2026 The Author(s). Journal of Modeling & Simulation in Electrical & Electronics Engineering, published by Semnan University Press. This is an open access article under the CC-BY 4.0 license. (<https://creativecommons.org/licenses/by/4.0/>)

speech and writing. In recent years, significant advances have been made in the diagnosis of Parkinson's disease through neurological markers such as EEG. EEG signal examination has advantages, including simple calculation, non-invasiveness, ease of access, low cost, etc. The physiological state of the brain can be inferred from EEG signals, and abnormalities can be detected by detecting unusual frequency patterns. Nowadays, machine learning methods are very efficient for feature extraction. The advantages of using machine learning models include improved prediction accuracy, improved efficiency and speed of data processing, pattern recognition, continuous improvement, no need for human intervention, and the possibility of examining large and complex data. In this paper, various methods are used to extract various features, including linear-nonlinear, time-frequency, etc. Various methods are then used to reduce and select features, and as a result, patients with PD are identified using effective classifiers. Usually, such a system involves extracting features from EEG signals and using classification techniques to classify signals into healthy and diseased classes.

II. METHODOLOGY

Parkinson's disease (PD) is a progressive neurological disorder. It is named after the researcher James Parkinson, who first described the motor condition in the 19th century, when it was known as the "shaking palsy". The main motor symptoms of PD include bradykinesia, resting tremor, and loss of postural reflexes. However, the condition is also associated with a wide range of non-motor symptoms. These non-motor symptoms have significant implications for daily functioning, personal care, and health-related costs. Movement disorders are primarily caused by a decrease in dopamine levels in the body, as a result of degeneration of the substantia nigra pars compacta. In fact, the appearance of motor symptoms may be a midpoint in the progression of the disease for many patients, before motor and non-motor symptoms worsen after a clear clinical diagnosis. Medications such as levodopa are often prescribed to manage some of the motor symptoms, especially for the problem of bradykinesia. However, there is no cure for Parkinson's disease. Therefore, symptom management is of primary importance in order to maintain health in these patients [1].

Cognitive decline is increasing in patients with PD. Therefore, a better understanding of factors associated with conversion and non-conversion is needed. However, there could be many reasons why participants may be classified as MCI-PD, including, medications, depression, infection, metabolic disorder, or cognitive function due to low educational attainment. MCI may be present in the form of frontal lobe damage in 50% of PD patients. Given the complexity of classifying MCI-PD, recent criteria from the Task Force on Movement Disorders (TF-MDS) attempted to exclude participants who had poor cognitive function due to comorbid conditions or severe emotional or psychiatric disturbances from being classified as MCI-PD [2].

The EEG is a method of recording electrical activity on the human scalp, which indicates the macroscopic activity of the subsurface layer of the brain [3]. One of the greatest advantages of this method is that it is non-invasive, and electrodes are placed along the scalp outside the human body. EEG measures

voltage changes produced by ion currents in the neurons of the brain. Clinically, it refers to the recording of spontaneous electrical activity of the brain over a period of time and is recorded by multiple electrodes [4].

The EEG signal is often used to detect lesions that cause abnormalities in EEG readings because they affect the normal functioning of the brain in a person with a neurological condition. This method can be used as a method of diagnosing tumors, strokes, and other brain disorders, even though this use has decreased with the advent of anatomical and high-resolution imaging techniques such as magnetic resonance imaging (MRI), it is still more difficult to record compared to the EEG signal [5].

The brain is the main source of input features, as EEG records brain activity through electrodes. The structure of the brain consists of three main parts: the forebrain, midbrain, and hindbrain, each of which is made up of several sub-parts [6].

The forebrain is the most important part because it is the largest part of the brain and is also known as the cerebral cortex. The cerebrum is the largest part of the human brain and is associated with higher brain functions, such as thinking and action. Gray matter is made up of nerve cells, and white matter is made up of nerve fibers that carry signals between nerve cells in other parts of the brain and the body. The white matter is then the actual part of the brain that produces the electrical signals recorded during an EEG test. The neuron is the basic cell of the brain. It has three main parts: dendrites, axon, and cell body, or soma. The dendrite is where the neuron receives input from other cells. (Fig. 1) As the dendrites extend, they develop leaf-like structures called spines. The output structure of the neuron is the axon. When a neuron wants to communicate with another neuron, it sends an action potential down the axon [7].

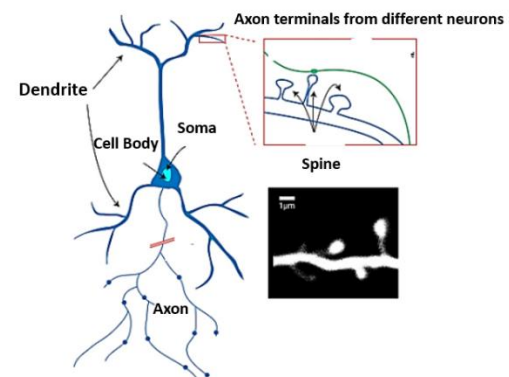


Fig. 1. Structure of neurons [25]

Neurons are essentially electrical devices, and their ability to carry voltage is guaranteed by their structure. The cell membrane, which is the boundary between the inside and outside of the cell, contains a large number of channels that allow the flow of positive and negative ions in both directions of the cell. The constant flow of positive and negative ions causes the membrane potential not to be static, but to constantly rise and fall depending on the inputs from other neurons [8].

When the membrane potentials of neurons reach a certain value, the action potential threshold, due to continuous excitatory and inhibitory activity, an action potential occurs, creating a threshold (Fig. 2).

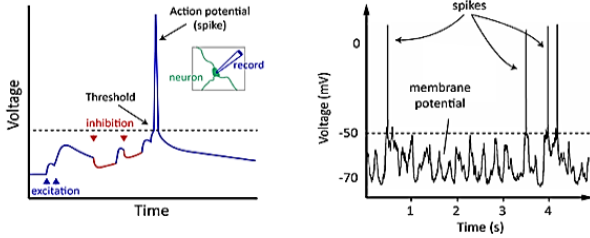


Fig. 2. Diagram of action potential jumps after the voltage reaches the action potential.

Resting-state electroencephalogram data is composite in nature and is broken down into medical frequency bands of delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (greater than 30 Hz). Each frequency band carries different physiological information related to brain activity. Delta and theta bands are revealed during the encoding of information and the formation of new memories. The alpha band represents information retrieval from memory and attention [9].

Beta and gamma bands are associated with perception, learning, and attention [10].

In patients with Parkinson's, the values obtained by each of these strips are distorted compared to a healthy patient. Resting-state activity in the brain is spontaneous but exhibits spatial and temporal organization. It is also useful in helping to map the internal connections in the brain to understand disease progression and ultimately help to understand underlying functional changes [11].

Machine learning is a subset of artificial intelligence that allows systems to learn from data and make predictions or decisions without explicit programming. This process typically involves the steps of data collection, data preprocessing, model training, model evaluation, and model deployment. Machine learning is divided into three main categories: supervised learning, in which data consists of specific inputs and outputs, and the goal is to make accurate predictions based on the inputs; unsupervised learning, in which data is unlabeled and outputs are specific, and the goal is to identify hidden patterns and structures in the data; and reinforcement learning, in which the agent interacts with the environment and learns how to act by receiving rewards or punishments. Machine learning is used in many areas, including medicine for diagnosing diseases and analyzing medical images, finance for fraud detection and market forecasting, marketing for analyzing customer behavior and recommenders, self-driving cars for identifying obstacles, and natural language processing (NLP) for machine translation.

Recently, machine learning (ML) has become a popular tool for diagnosing medical diseases due to its ease of implementation and high accuracy. Dysphonia measurement was proposed as a way to differentiate between people with PD and healthy individuals by analyzing changes in voice frequency [12].

The popularity of convolutional neural networks (CNN) has increased significantly, thanks to their impressive performance in various tasks such as image segmentation, object tracking, and classification. Neural processing techniques and methods are inspired by the structures of the human brain [13].

The convolution layer is the main element of CNN used for image recognition. Convolutional layers are important because they can learn to identify and extract important features from images without the need for manually engineered feature extraction. One of the layers in a convolutional neural network involves passing the extracted feature set to a fusion layer, which can perform operations such as maximum or median fusion, depending on the application. The fusion layer helps to compress the image while preserving essential information [14] (Fig. 3).

The main function of the connected layer is to establish connections between all neurons in the previous layer and each neuron in the next layer, facilitating the classification of features obtained from other layers.

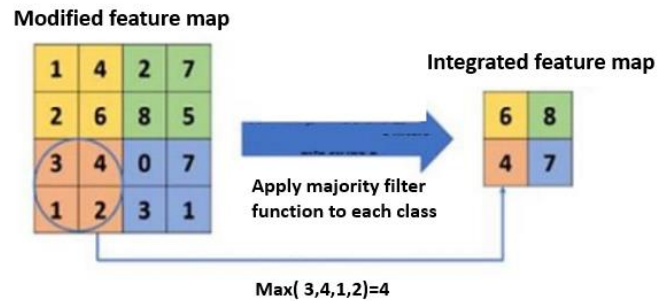


Fig. 3. The method in which maximum integration is displayed

Activation functions enable nonlinear transformations in deep neural networks and make it easier to learn complex tasks during the training process. The hyperbolic tangent function is commonly used as a nonlinear activation function when creating multiple layers in a perceptron neural network, and its output range is usually in the interval $(-1, 1)$ as shown in (1).

$$f(x) = \frac{2}{1+e^{-2x}} - 1 \quad (1)$$

A newer solution has emerged to replace deep architectures, and one of these solutions is modified linear units, as shown in (2).

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (2)$$

Compared to traditional units, modified linear units offer several advantages, such as faster computations and improved gradient propagation due to their lack of saturation. The loss function is a useful tool for evaluating the effectiveness of a network model in handling labeled data. Two common types of loss functions are cross entropy and mean square error (MSE), each of which is calculated differently and can help with gradient optimization [15].

Evaluation is a key aspect of any experiment, as it determines the accuracy of the classification for a given dataset. Metrics such as accuracy (based on the confusion matrix), recall, precision, etc., can be used. A hash matrix is a tool used to represent and improve the accuracy of classification models. This matrix contains columns that represent the samples in the predictor species [16].

III. RELATED WORKS

Bibbo et al. [17] presented a new approach to detect emotional health using the SNN model. Today, artificial intelligence identifies types of activities by analyzing data from different sources. Human facial expressions change with different health states. SNN networks make it possible to create a model with a reduced number of levels compared to existing models. As a result, the accuracy provided is reported to be 98.4% with a learning rate of 0.0001.

Abdullah et al. [18] investigated and published a method for diagnosing Parkinson's disease based on deep transfer learning using optimal feature selection. Today, handwritten records are one of the dominant mechanisms used for diagnosing and assessing PD. This reference proposes an efficient deep learning model that can help in the early diagnosis of Parkinson's disease. They reported a detection accuracy of 98% and an area under the curve of 90%. Rehman et al. [19] investigated a method for diagnosing Parkinson's disease using a hybrid LSTM-GRU deep learning model. In this study, a dataset consisting of 31 male and female patients with 195 different voices was collected and used.

Li et al. [20] investigated a method for diagnosing and classifying Parkinson's disease using EEG signals based on a deep CNN-LSTM model. They presented two hybrid deep neural networks (DNNs) that combine convolutional neural networks with short-term memory for the diagnosis of Parkinson's disease based on EEG signals. In that study, specificity of 97.6%, sensitivity of 97.1%, and accuracy of 98.6% were reported for a parallel model, and specificity of 99%, sensitivity of 98.5%, and accuracy of 99.7% were reported for a serial model. Tolu et al. [21] have explored an important perspective for investigating dementia diseases with neurobotic approaches. Recently, advances in this field have made it possible to test some neuroscience hypotheses related to brain diseases. However, the lack of biological plausibility of developed brain models and musculoskeletal systems has limited understanding of the brain mechanisms that lead to deficits in motor and cognitive tasks.

Matin et al. [22] conducted a review study based on machine learning approaches for diagnosing Parkinson's disease. In this review, machine learning (ML) methods were used to analyze PD diagnosis through resting-state EEG motor activation tests.

Al-Jalal et al. [23] investigated a method for diagnosing Parkinson's disease from resting-state EEG signals using joint spatial pattern, entropy, and machine learning methods. In this study, joint spatial pattern-based approaches are proposed for diagnosing Parkinson's disease in two cases, without and with medication.

Loh et al. [24] proposed a method using the Gabor transform and a deep neural network for signal-based diagnosis of Parkinson's disease. In that study, a deep learning model was proposed for automatic PD diagnosis. They used EEG signals from 16 healthy individuals and 15 PD patients for analysis.

Alzubaidi et al. [25] presented a review method to investigate the role of neural networks in the diagnosis of Parkinson's disease. This study aimed to review and summarize the applications of neural networks for the diagnosis of PD. Two medical databases (PubMed and IEEE) were used to identify relevant studies. Loh et al. [26] presented a review based on

deep learning models for the automatic detection of Parkinson's disease. Computer-aided diagnostic (CAD) tools, based on artificial intelligence methods, can automatically detect PD.

Geraedts et al. [27] presented a machine learning method for automated EEG signal-based biomarkers of cognitive impairment during deep brain stimulation screening in patients with Parkinson's disease. They used a fully automated method for 112 patients with PD that predicted the EEG signal as input and class labels as output.

IV. PROPOSED METHOD

EEG signals are a non-invasive and efficient tool that provides valuable information about the electrical activity of the brain and can help identify specific patterns associated with cognitive disorders. Fig. 4 shows a schematic diagram of the proposed method and depicts the different stages of the research from data collection to the final analysis and interpretation of the results. This method can help develop new diagnostic tools and improve treatment processes for patients with Parkinson's disease and related cognitive disorders. The proposed method consists of five main steps: database acquisition, preprocessing, feature extraction, feature selection, and classification. First, a suitable database containing EEG signals from individuals with Parkinson's disease is collected. In the next step, the acquired signals are preprocessed by filtering and segmentation. Next, high-level features are extracted from these signals, and then their dimensions are reduced by four different methods. In the last step, Parkinson's disease is classified and diagnosed by different machine learning methods.

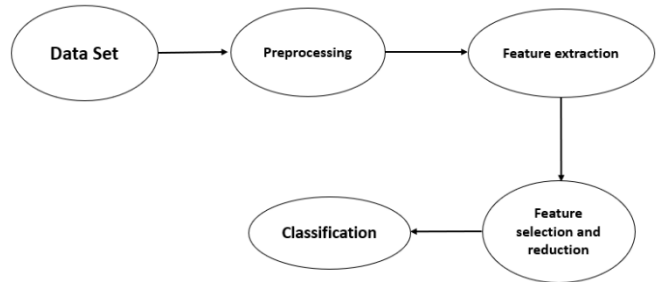


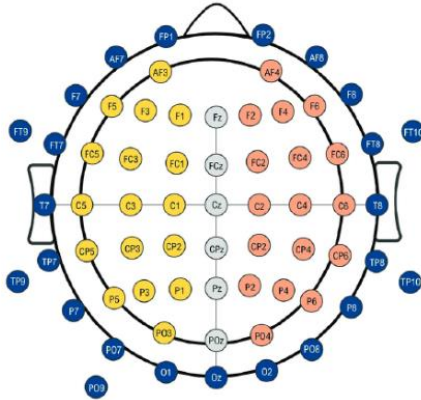
Fig. 4. General scheme of the proposed method

To simulate the proposed method, the PD Interval Timing database available on the University of New Mexico (UNM) portal was used. [15] Also, with the aim of reviewing and summarizing the applications of neural networks for PD diagnosis and to identify relevant studies, two medical databases (PubMed and IEEE) were used. Then, the presented approach was adopted to combine the extracted data. This information is fully substantiated and reliable and is used in all simulations and studies.

This database is a collection of EEG signals from 37 healthy volunteers (23 males and 18 females) and 92 patients with PD (56 males and 24 females). The average age of the volunteers in the healthy group is 71.3 years, and in the Parkinson's group is 68.7 years.

Clinical information of the patients, along with the results of the Montreal Cognitive Assessment (MOCA) test were obtained for the patients in two groups of healthy and sick. The available EEG signals were recorded with a sampling frequency

of 500 Hz and as 63 channels. Fig. 5 shows the electrodes used in signal recording and their locations. The EEG signal is a very complex signal that usually requires preprocessing for analysis. The preprocessing steps performed are as follows:



complex data, such as (EEG signals. In this method, the main goal is to identify and extract common factors among different features of the EEG signal.

B. *t*-Stochastic Neighbor Embedding

This is a method for reducing the dimensionality of data that is used in the analysis of complex, multidimensional data, such as image data or biological signals. This method is especially effective for visualizing data in low dimensions (usually 2 or 3 dimensions). This method is particularly effective for visualizing data in low dimensions (2D or 3D) and can help identify hidden patterns and relationships.

C. Principal Component Analysis (PCA)

It is a statistical technique used to reduce data dimensions and extract key features from large data sets. In the steps of performing PCA, the data must first be standardized. This is usually done using the Z-score transformation. After standardization, the next step is to calculate the covariance matrix. Then, the eigenvalues and eigenvectors are calculated from the covariance matrix. This method can also help reduce noise in the data. The PCA method is used in various fields, including machine vision, where it is used to reduce the dimensions of images and extract key features.

After extracting features from EEG signals, the next step in data analysis is machine learning-based classification. This step is performed to identify patterns in the data and diagnose diseases, especially Parkinson's disease. For this purpose, first, the feature matrix, which contains the values extracted from EEG signals, is partitioned using the k-fold cross-validation method.

In this partitioning, 80% of the data was used for training and 20% for testing across the five iterations. After segmenting the data, several different classification methods have been used to identify and diagnose Parkinson's disease. One of these methods is the Support Vector Machine (SVM), which is implemented with three different types of kernels (such as linear, polynomial, and Radial Gaussian Basis Function (RBF). Choosing the right kernel for SVM can have a great impact on classification accuracy, as kernels help the model better separate data in the nonlinear feature space.

To accurately evaluate the models, several evaluation indicators have been used. These indicators include accuracy, sensitivity, and specificity. Accuracy represents the ratio of correct predictions to total predictions. Sensitivity describes how many of all true positives are correctly identified. This indicator is very important for diagnosing diseases such as Parkinson's. Specificity indicates how many of all true negative samples are correctly identified.

The McNemar test is used, especially in medical research and studies, to measure the accuracy of the performance of a drug or a specific type of patient, and it determines the sensitivity and specificity of the measurement. In this way, the accuracy of a test can be calculated. The McNemar test is performed between two groups of patients and healthy people to determine the correct and incorrect results of the test. The diagnosis of whether a person has a disease or not is made by a specialist or clinical symptoms, which is definitely correct. However, since specialists are not available everywhere, in this test, the presence or absence of the disease is diagnosed in the individual through a chemical test. Since this test has two responses or two-state values

(sick or healthy) and there are also two types of tests (chemical test results and expert opinion), the prerequisites for implementing the McNemar test are a two-by-two agreement table. In the McNemar test, consider a contingency table like Table I. In the first column (expert opinion), the situation for test number 1 is written, and in the second column, the situation for test number 2 (chemical test).

TABLE I
Agreement Table and Medical Test Error in the McNemar Test

Fact	Test result = patient	Test result = healthy	Row sum
Patient	a	b	a+b
Healthy	c	d	c+d
Column sum	a+c	b+d	n

Here, accuracy in a medical test refers to its ability to correctly distinguish between a sick person and a healthy person, and is determined by (5).

$$\text{Accuracy} = \frac{a+d}{a+c+b+d} \quad (5)$$

In this case, sensitivity is measured by the following equation.

$$\text{sensitivity} = \frac{a}{a+c} \quad (6)$$

On the other hand, the Specificity, which is a kind of accuracy of a medical test for finding non-diseased people, is obtained by the following equation.

$$\text{specificity} = \frac{d}{d+b} \quad (7)$$

V. SIMULATION RESULTS

Open-access EEG signals from 37 people with Parkinson's disease were used. EEG signals have different durations for different people. Fig. 6 shows a sample of a 1-minute EEG signal for three randomly selected subjects from the database.

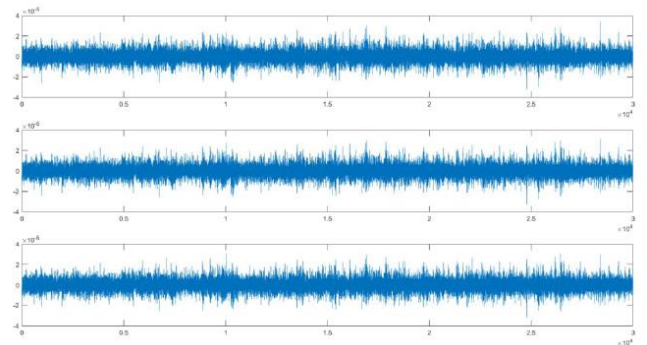


Fig. 6. Example of EEG signals from the database of three subjects

Fig. 7 illustrates a comparison of Filtered EEG signals obtained from healthy subjects and PD patients.

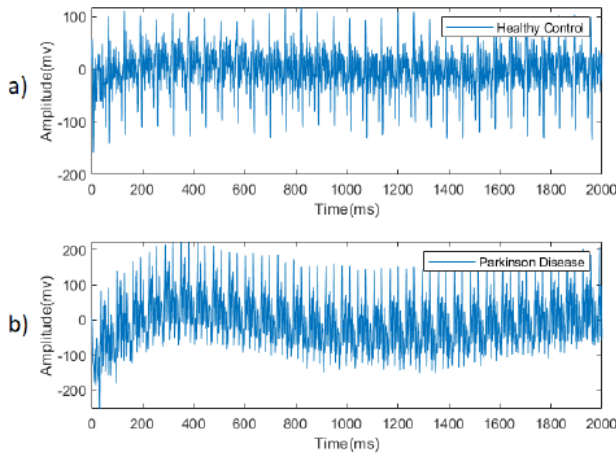


Fig.7. Comparison of EEG Signals for Healthy Subjects and Parkinson's Disease Patients. a) healthy control b) Parkinson's disease.

Four classification methods were used, including SVM, K-Nearest Neighbor (KNN), Decision Tree (DT), and Random Forest (RF). SVM can classify nonlinear data well and is usually suitable for high-dimensional classification problems. KNN simply classifies data based on its nearest neighbors and is easy to understand. DT also helps in analyzing data hierarchically and can be easily interpreted. The performance of the models was evaluated using accuracy, sensitivity, and specificity indices. Random forest, or random decision forest, is an ensemble learning method for classification and regression, which is based on a structure consisting of a large number of decision trees, working on the training time and output of classes (classification) or for predictions of each tree separately. Random forests are suitable for decision trees that suffer from overfitting in the training set. The performance of random forests is usually better than decision trees, but this performance improvement also depends to some extent on the type of data.

A. Results of Feature Reduction

Feature Reduction is a multivariate method used to reduce dimensionality and identify hidden structures in complex data, especially EEG signals. This method helps to identify common factors among different features of the EEG signal and allows the creation of new variables called features that explain the variations in the data. The main goal of factor analysis in EEG signal processing is to reduce the dimensionality of the data and simplify the analyses. Table II shows the performance of this method. This table also shows the neural network training time and the memory usage in kilobytes (KB).

TABLE II
Average Results of the 5-fold Feature Reduction Method

Model	Accuracy	Sensitivity	Specificity	Training Time	memory usage
RBF	81.01	88.20	72.65	6.54 s	13.3 KB
KNN	93.41	98.99	93.41	5.45 s	17.2 KB
DT	91.30	93.55	90.47	5.85 s	18.2 KB
RF	86.3	83.7	84.3	4.24 s	19.2 KB

B. Results of t-Stochastic Neighbor Embedding

t-SNE is a multidimensional data analysis algorithm used to visualize data in a reduced-dimensional space (usually two-dimensional or three-dimensional). Using t-SNE, complex, multidimensional data can be represented in a uniform and regular manner in a reduced space. The algorithm creates a new space in which similar data are close together and dissimilar data are far apart. In other words, data that were close together in the original space remain close together in the reduced space, and data that were far apart in the original space are far apart in the reduced space. t-SNE is well-suited for visualizing data that is represented nonlinearly in multidimensional space. Overall, t-SNE is an efficient and useful algorithm for data visualization that is widely used in various fields such as data science, machine vision, statistics, genomics, cybersecurity, natural language processing, music analysis, cancer research, bioinformatics, geology interpretation, and medical signal processing. The t-SNE method is an effective technique for data dimensionality reduction, especially useful in the analysis of complex and multidimensional data, such as image data and biological signals. t-SNE maps nearby points in the original space to nearby points in the new space, while preserving the distances between more distant points. The operation of t-SNE includes the steps of calculating the distances between points, calculating the probability that each point is close to other points, and mapping the points to a lower-dimensional space using the t-distribution. This method preserves the local structure of the data well and can identify natural clusters.

Table III shows the performance of this method.

TABLE III
Average Results of the 5-fold t-SNE Method

Model	Accuracy	Sensitivity	Specificity	Training Time	memory usage
RBF	80.99	85.97	75.57	6.54 s	12.9 KB
KNN	96.80	99.00	92.42	5.45 s	17.8 KB
DT	90.68	93.87	87.79	5.85 s	17.9 KB
RF	81.2	84.7	88.1	4.55 s	18.8 KB

C. Results of Principal Component Analysis (PCA)

PCA is an effective statistical technique for reducing the dimensionality of data and extracting key features from large data sets. This method helps to identify hidden patterns and relationships. The steps of performing PCA include standardizing the data using Z-score transformation, calculating a covariance matrix to examine the correlation between variables, and extracting eigenvalues and eigenvectors from this matrix. The principal components with larger eigenvalues are selected, and the data are transformed into a new space with lower dimensions. The advantages of PCA include reducing dimensionality and discovering patterns, but it also has disadvantages, including the possibility of losing important information and difficulty in interpreting the components. This method is used in various fields such as machine vision and helps in the effective analysis of large and complex data. Table IV shows the performance of this method.

TABLE IV
Average Results of the 5-fold PCA Method

Model	Accuracy	Sensitivity	Specificity	Training Time	memory usage
RBF	70.28	72.97	73.89	5.61 s	12.2 KB
KNN	82.80	86.90	81.39	4.75 s	15.4 KB
DT	70.28	71.29	67.00	4.95 s	16.2 KB
RF	85.2	86.5	81.5	4.25 s	17.6 KB

In this regard, the t-SNE method is an effective and advanced technique compared to the other two methods. The t-SNE method performs better, especially in analyzing complex and multidimensional data, such as image data and biological signals. Unlike the PCA method, which operates on the overall variance of the data and may ignore local structure, t-SNE specifically focuses on preserving local relationships. Also, feature reduction that seeks to identify common factors may not be able to capture complex and nonlinear structures well. While t-SNE effectively maps points to a lower-dimensional space using the t-distribution, its ability to preserve local structure and identify clusters makes it a suitable option for analyzing complex data. In Table II, the performance of this method is examined in detail, and the results show the superiority of t-SNE over PCA and feature reduction in certain areas of data analysis. Finally, due to the unique features of t-SNE and its ability to preserve the local structure of the data, this method is recognized as a powerful tool in complex data analysis, especially in research and applied fields. According to the results obtained, it can be seen that the KNN method has reported the best performance with accuracy, sensitivity, and specificity (96.80, 99.00, and 92.42) among all 3 classifiers. Table V shows the comparison of the proposed work with previous research. According to the results, it can be seen that the proposed method has better performance. Compared to other classifiers, the KNN method achieved the best performance with an accuracy of 96.80, sensitivity of 99.00, and specificity of 92.42. These results show that KNN acts as a strong option for data classification and can be used in various applications that require high accuracy.

VI. CONCLUSION

This paper examines and analyzes EEG signals from patients with Parkinson's disease and compares them with signals from healthy individuals. Parkinson's disease is a progressive neurological disorder that has significant effects on patients' movements and daily activities due to the destruction of dopamine-producing nerve cells in the brain. The main objective of this paper is to identify specific patterns in the EEG signals of patients with PD and to better understand the electrical changes of the brain compared to healthy individuals. This study can help develop new methods in the diagnosis and management of Parkinson's disease. This database contains EEG signals from 37 healthy volunteers and 92 patients with PD. The average age of the volunteers in the healthy group was reported to be 71.3 years, and in the patient group, to be 68.7 years. This age information could potentially influence the results of the analyses and help to better understand the differences between the two groups. EEG signals were recorded with a sampling frequency of 500 Hz and 63 channels. This number of channels allows researchers to examine the electrical

changes in the brain from different angles. To increase the quality of signal processing, pre-processing steps including filtering, noise removal, normalization, and segmentation were carefully performed. In the filtering stage, a band-pass filter with a frequency range of 1 to 30 Hz is used to remove additional noise. This filter helps separate the desired signals from unnecessary noise and increases the quality of the data. After filtering, the signals were normalized using the z-score normalization method and divided into ten-second segments. Next, various features of EEG signals were extracted, and for further analysis, feature reduction methods such as feature reduction, t-SNE, and PCA were used. After feature extraction, machine learning-based classification was performed. For this purpose, the data was divided into two groups, training and testing, using 5-fold cross-validation.

TABLE V
Comparison of the Proposed Method with other Methods

Model	Feature	Classification	Accuracy
[18]	Deep Feature	Deep learning models	96.45
[23]	A variety of different features	KNN-SVM	95
[24]	A variety of different features	KNN	91.70
Proposed method	A variety of different features	KNN	96.80

Accuracy refers to the ratio of the number of correct predictions to the total number of predictions, while sensitivity indicates the model's ability to identify patients with Parkinson's disease, and specificity indicates the model's ability to identify healthy individuals. The results of this paper could help improve diagnostic processes and identify abnormal movement patterns in patients with PD, and lead to the development of new methods for diagnosing and managing this disease. This research can help develop optimal treatment plans and interventions for patients with PD and lead to improved quality of life. In general, KNN is considered a suitable option for data analysis due to its simplicity and efficiency.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support for this work that was provided by the Islamic Azad University of Shoushtar Branch.

FUNDING STATEMENT

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.

CONFLICTS OF INTEREST

The author declares that there is no conflict of interest regarding the publication of this article.

AUTHORS CONTRIBUTION STATEMENT

Seyed Ali Sadatnoori: Conceptualization; Numerical Implementation; Experimental Validation; Writing of the Manuscript,

Hakim Golshahi: Data analysis and interpretation; Conceptualization and study design.

AI DISCLOSURE

AI cannot be listed as an author. Only human authors take responsibility.

DATA AVAILABILITY

All data generated or analyzed during this study are included in this published article.

REFERENCES

- [1] S. M. Park, B. Jeong, D. Y. Oh, C. H. Choi, H. Y. Jung, J. Y. Lee, and J. S. Choi, "Identification of major psychiatric disorders from resting-state electroencephalography using a machine learning approach," *Front. Psychiatry*, vol. 12, Art. no. 707581, 2021.
- [2] S. P. Kim, "Preprocessing of EEG," in *Computational EEG Analysis: Methods and Applications*, pp. 15–33, 2018.
- [3] S. Fahn and D. Sulzer, "Neurodegeneration and neuroprotection in Parkinson disease," *NeuroRx*, vol. 1, no. 1, pp. 139–154, 2004.
- [4] M. Z. Lubis, "Signal processing for power spectral density (PSD)," in *Signal Processing for Marine Acoustic and Dolphin Using MATLAB*, [publication details not provided], 2016.
- [5] P. Cunningham and S. J. Delany, "k-nearest neighbor classifiers (with Python examples)," *arXiv preprint arXiv:2004.04523*, 2020.
- [6] E. van Diessen, T. Numan, E. van Dellen, A. W. van der Kooij, M. Boersma, D. Hofman, R. van Lutterveld, B. W. van Dijk, E. C. W. van Straaten, A. Hillebrand, and C. J. Stam, "Opportunities and methodological challenges in EEG and MEG resting state functional brain network research," *Clin. Neurophysiol.*, vol. 126, no. 8, pp. 1468–1481, 2015.
- [7] G. J. Kress and S. Mennerick, "Action potential initiation and propagation: Upstream influences on neurotransmission," *Neuroscience*, vol. 158, no. 1, pp. 211–222, 2009.
- [8] M. Azhari, A. Alaoui, Z. Achraoui, B. Ettaki, and J. Zerouaoui, "Adaptation of the random forest method: Solving the problem of pulsar search," in *Proc. 4th Int. Conf. Smart City Applications*, Oct. 2019, pp. 1–6.
- [9] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 1, pp. 20–28, 2021.
- [10] K. Sidiropoulou, E. K. Pissadaki, and P. Poirazi, "Inside the brain of a neuron," *EMBO Rep.*, vol. 7, no. 9, pp. 886–892, 2006.
- [11] J. S. Almeida, P. P. Rebouças Filho, T. Carneiro, W. Wei, R. Damaševičius, R. Maskeliūnas, and V. H. C. de Albuquerque, "Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques," *Pattern Recognit. Lett.*, vol. 125, pp. 55–62, 2019.
- [12] M. M. Mijwil, I. E. Salem, and M. M. Ismaeel, "The significance of machine learning and deep learning techniques in cybersecurity: A comprehensive review," *Iraqi J. Comput. Sci. Math.*, vol. 4, no. 1, p. 10, 2023.
- [13] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.
- [14] Q. Chen, Q. Xie, Q. Yuan, H. Huang, and Y. Li, "Research on a real-time monitoring method for the wear state of a tool based on a convolutional bidirectional LSTM model," *Symmetry*, vol. 11, no. 10, Art. no. 1233, 2019.
- [15] A. Singh, R. C. Cole, A. I. Espinoza, A. Evans, S. Cao, J. F. Cavanagh, and N. S. Narayanan, "Timing variability and midfrontal ~4 Hz rhythms correlate with cognition in Parkinson's disease," *npj Parkinson's Dis.*, vol. 7, no. 1, Art. no. 14, 2021.
- [16] M. B. Er, E. Isik, and I. Isik, "Parkinson's detection based on combined CNN and LSTM using enhanced speech signals with variational mode decomposition," *Biomed. Signal Process. Control*, vol. 70, Art. no. 103006, 2021.

- [17] L. Bibbo, F. Cotroneo, and M. Vellasco, "Emotional health detection in HAR: New approach using ensemble SNN," *Appl. Sci.*, vol. 13, no. 5, Art. no. 3259, 2023.
- [18] S. M. Abdullah, T. Abbas, M. H. Bashir, I. A. Khaja, M. Ahmad, N. F. Soliman, and W. El-Shafai, "Deep transfer learning based Parkinson's disease detection using optimized feature selection," *IEEE Access*, vol. 11, pp. 3511–3524, 2023.
- [19] A. Rehman, T. Saba, M. Mujahid, F. S. Alamri, and N. El Hakim, "Parkinson's disease detection using hybrid LSTM-GRU deep learning model," *Electronics*, vol. 12, no. 13, Art. no. 2856, 2023.
- [20] K. Li, B. Ao, X. Wu, Q. Wen, E. Ul Haq, and J. Yin, "Parkinson's disease detection and classification using EEG based on deep CNN-LSTM model," *Biotechnol. Genet. Eng. Rev.*, vol. 40, no. 3, pp. 2577–2596, 2024.
- [21] S. Tolu, B. Strohmmer, and O. Zahra, "Perspective on investigation of neurodegenerative diseases with neurorobotics approaches," *Neuromorphic Comput. Eng.*, vol. 3, no. 1, Art. no. 013001, 2023.
- [22] A. M. Maitin, J. P. Romero Muñoz, and A. J. García-Tejedor, "Survey of machine learning techniques in the analysis of EEG signals for Parkinson's disease: A systematic review," *Appl. Sci.*, vol. 12, no. 14, Art. no. 6967, 2022.
- [23] M. Aljalal, S. A. Aldosari, K. Al Sharabi, A. M. Abdurraqueeb, and F. A. Alturki, "Parkinson's disease detection from resting-state EEG signals using common spatial pattern, entropy, and machine learning techniques," *Diagnostics*, vol. 12, no. 5, Art. no. 1033, 2022.
- [24] H. W. Loh, C. P. Ooi, E. Palmer, P. D. Barua, S. Dogan, T. Tuncer, M. Baygin, and U. R. Acharya, "GaborPDNet: Gabor transformation and deep neural network for Parkinson's disease detection using EEG signals," *Electronics*, vol. 10, no. 14, Art. no. 1740, 2021.
- [25] M. S. Alzabadi, U. Shah, H. Dhia Zubaydi, K. Dolaat, A. A. Abd-Alrazaq, A. Ahmed, and M. Househ, "The role of neural network for the detection of Parkinson's disease: A scoping review," *Healthcare*, vol. 9, no. 6, Art. no. 740, Jun. 2021.
- [26] H. W. Loh, W. Hong, C. P. Ooi, S. Chakraborty, P. D. Barua, R. C. Deo, J. Soar, E. E. Palmer, and U. R. Acharya, "Application of deep learning models for automated identification of Parkinson's disease: A review (2011–2021)," *Sensors*, vol. 21, no. 21, Art. no. 7034, 2021.
- [27] V. J. Graedts, M. Koch, M. F. Contarino, H. A. M. Middelkoop, H. Wang, J. J. van Hilten, T. H. W. Bäck, and M. R. Tannemaat, "Machine learning for automated EEG-based biomarkers of cognitive impairment during deep brain stimulation screening in patients with Parkinson's disease," *Clin. Neurophysiol.*, vol. 132, no. 5, pp. 1041–1048, 2021.

BIOGRAPHIES

Seyed Ali Sadatnoori received his BSc degree in Electrical Engineering from the Islamic Azad University of Dezful Branch and MSc degree in Electrical Engineering from the Central Tehran branch of Islamic Azad University, respectively, in 2001 and 2006. He received the PhD degree in Electrical Engineering from the Tehran Science and Research Branch of Islamic Azad University in 2016. He is also an Assistant Professor at the Shoushtar Branch of Islamic Azad University. His current research interests include mobile ad hoc networks, wireless networks, transmitter and receiver circuits, analog-to-digital modulators, integrated circuits, sigma-delta modulators, and neural networks.

Hakim Golshahi received his BSc degree in Applied Physics from Kharazmi University in 2002, his MSc degree in Physical Oceanography from Tarbiat Modares University in 2006, and his PhD degree in Physical Oceanography from the Science and Research Branch of Islamic Azad University in 2011. He is currently an Assistant Professor in the Department of Physics at the Shoushtar Branch of Islamic Azad University. His research interests include numerical simulation of oceanographic and physical phenomena using mathematical and numerical methods.