RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF
MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt2456

Comprehensive EEG Signal Feature Extraction for Neurological Disorder Diagnosis: Focus on Alzheimer's, Parkinson's, and Seizure Disorders

K. Nanthini [a], D. Sivabalaselvamani [b,*], M.C. Madhan Kumar [c,*], R. Kaviya [c]

[a] Department of Computer Science & IT, Jain University, Jain Nagar, Bengaluru, Karnataka, India.
[b] School of Information Science, Presidency University, Bengaluru, Karnataka, India.
[c] Department of Computer Applications, Kongu Engineering College, Perundurai, Erode, Tamil Nadu, India
* Corresponding Author Email: sivabalaselvamani@gmail.com
DOI: https://doi.org/10.54392/irjmt2456
Received: 02-02-2024; Revised: 29-08-2024; Accepted: 11-09-2024; Published: 20-09-2024

Abstract: This research paper examines the use of Electroencephalogram (EEG) signal feature extraction for diagnosing neurological disorders, specifically Alzheimer's, Parkinson's, and seizure disorders. It evaluates various methods for categorizing EEG signals, including time-domain, frequency-domain, and statistical transformations emphasizing their effectiveness in distinguishing relevant brainwave patterns (beta, alpha, theta, delta) from artifacts like eye blinks and muscle movements. The study highlights the challenges in artifact removal and provides an overview of key feature extraction techniques, particularly in the time and frequency domains. The implementation section details the application of machine learning algorithms to classify mental states using statistical features from EEG signals. The research identifies specific EEG patterns associated with Alzheimer's, Parkinson's, and seizure disorders, noting alterations in alpha, theta, and delta waves. The paper underscores the critical role of EEG feature extraction in diagnosing neurological disorders and recommends incorporating additional frequency-based methods to enhance predictive accuracy in future research.

Keywords: Electroencephalogram (EEG), Signal processing, Neurological disorders, Feature Extraction, Machine learning, Time-frequency methods

1. Introduction

Electroencephalography (EEG) has long been used to monitor brain activity, providing essential insights into various neurological disorders. Its non-invasive nature makes it a valuable tool in both research and clinical settings. EEG signals, consisting of multiple frequency bands (delta, theta, alpha, beta, and gamma), provide an in-depth understanding of brain functions and abnormalities, making it indispensable in diagnosing conditions such as Alzheimer's, Parkinson's, and seizure disorders. However, the complex and dynamic nature of EEG data makes it difficult to interpret, requiring robust feature extraction techniques for efficient analysis. Several advanced methods for EEG feature extraction have been developed to address these challenges. For instance, Discrete Wavelet Transform (DWT) for EEG feature extraction, revealing its effectiveness in identifying patterns associated with epilepsy [1]. Similarly, researchers have adopted techniques such as Fourier Transform and Empirical Mode Decomposition to decompose EEG signals into distinct components for better analysis [2]. Another notable contribution is the work of developing an automatic EEG classification system using machine learning algorithms [3].

One of the most critical hurdles in EEG signal analysis is the presence of noise and artifacts, such as eye blinks and muscle movements, which significantly affect the quality of data. Methods such as Independent Component Analysis (ICA) and Principal Component Analysis (PCA) have been widely used to mitigate these artifacts, with considerable success [4]. Additionally, advances in machine learning and signal processing have led to the development of novel feature extraction techniques, improving the accuracy and reliability of EEG-based diagnosis [5]. The high dimensionality and redundancy of EEG data often lead to computational inefficiencies and can result in model overfitting. Strategies like feature selection and dimensionality reduction, combined with cross-validation techniques, have been employed to overcome these issues [6]. Furthermore, advancements in deep learning architectures, such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, have shown promise in capturing complex temporal dependencies within EEG data [7]. Despite the significant progress in EEG-based neurological diagnosis, challenges remain, including the inter- and intra-subject variability in EEG signals and the limited availability of comprehensive datasets. Addresses the variability issue by introducing robust feature extraction techniques that improve the generalization capabilities of EEG-based diagnostic models [8]. Nonetheless, the use of EEG as a diagnostic tool for neurological disorders continues to grow, offering great potential for early detection and intervention.

2. Literature Review

2.1 Neurological Insights through EEG Analysis

EEG has been a widely used tool for diagnosing and monitoring neurological disorders. In Alzheimer's disease (AD), EEG patterns reveal a decline in alpha and beta oscillations, reflecting the deterioration in cortical networks and cognitive functioning. By extracting alpha and theta waves from EEG signals, early detection of Alzheimer's can be enhanced [9]. Similarly, theta rhythms have been associated with memory impairment, making them a critical feature for the diagnosis of AD [10]. EEG has also been pivotal in Parkinson's disease (PD) research. A study found that Parkinson's patients exhibit alterations in beta band activity, which is closely linked to motor dysfunction [11]. Additionally, gamma band abnormalities have been reported in PD patients, correlating with cognitive decline and motor symptoms [12].

2.2 Imaging Modalities in Cognitive Disorders

The comparison between EEG and other imaging modalities like functional magnetic resonance imaging (fMRI) has been extensively explored. While fMRI provides high spatial resolution, EEG is superior in terms of temporal resolution, making it ideal for real-time brain monitoring. For instance, studies have shown that combining EEG with fMRI offers a more comprehensive understanding of brain dynamics, especially in the context of cognitive disorders [13].

2.3 Cutting-Edge Approaches in Epilepsy Studies

Epilepsy research has greatly benefited from advances in EEG feature extraction techniques. For example, Wavelet Packet Decomposition (WPD) and Support Vector Machine (SVM) classifiers to achieve high accuracy in distinguishing epileptic seizures from normal brain activity [14]. Another study applied deep learning models to EEG signals, significantly improving seizure detection rates [15]. These methods have shown promise in reducing false positives in epilepsy detection, contributing to better patient outcomes.

2.4 Innovations in Parkinson's Disease and Seizure Diagnoses

EEG's application in diagnosing Parkinson's and seizure disorders has expanded in recent years. Highlighted how automated EEG analysis could be used to detect Parkinson's disease in its early stages by identifying specific changes in alpha and beta wave patterns [16]. For seizure disorders, utilized a combination of statistical features and machine learning algorithms to develop a highly accurate seizure prediction model based on EEG signals [17]. These approaches demonstrate the utility of EEG in clinical settings, especially in detecting and predicting neurological events.

3. Feature Extraction in Signal Processing: Unlocking Information through Machine Learning

Feature extraction is an important process in signal processing, where raw data is transformed into numerical features to enable meaningful analysis. Machine learning and deep learning algorithms play a crucial role in this process, especially for time series data and signals. Training these algorithms directly with raw signals can result in suboptimal outcomes due to high data velocity and information redundancy. Feature extraction methods are categorized into four types: time domain, frequency domain, time-frequency domain, and nonlinear methods. Frequency domain methods such as power spectral analysis provide insights into the frequency content of EEG signals. Time domain approaches, including Linear Prediction and Component Analysis, extract parameters based on time, bridging the gap between physical interpretation and spectral analysis. Time-frequency methods, such as Wavelet and Hilbert-Huang Transforms, capture transient features in both time and frequency dimensions. Nonlinear EEG analysis introduces parameters like Lyapunov Exponents and entropies, addressing the complexity of non-stationary signals and contributing to pioneering work in nonlinear dynamics. This paper focuses on the time and frequency domain aspects of EEG signals, employing machine learning algorithms for a comprehensive analysis. The objective is to demonstrate the effectiveness of feature extraction methods in enhancing the interpretability and utility of EEG data, contributing significantly to advancements in the diagnosis of neurological disorders and related research.

3.1 Time Domain

Time-domain analysis is a crucial approach for unraveling the temporal intricacies of EEG signals, offering valuable insights into the dynamic nature of brain activity over time. The process involves extracting key features such as Amplitude, Duration, Slope, and Waveform Patterns, providing a comprehensive understanding of how brain signals evolve temporally. This analytical method plays a pivotal role in decoding the temporal characteristics inherent in EEG data. Among the key features extracted through time-domain analysis, the First Difference represents the discrete variance between consecutive EEG signal values. The Normalized First Difference adjusts for amplitude variations, enhancing the precision of the analysis. The Second Difference introduces an additional layer by considering the difference of the first difference, while the Normalized Second Difference refines this by normalizing for amplitude changes. Metrics like Mean Curve Length and Mean Energy capture overall signal characteristics, while Skewness and Kurtosis delve into distribution properties. Parameters like Hjorth Activity, Hjorth Mobility, and Hjorth Complexity offer insights into signal energy, frequency variations, and overall signal complexity, respectively. Together, these features form a comprehensive toolkit for understanding the temporal dynamics of EEG signals.

3.2 Frequency Domain

When analyzing EEG signals in the frequency domain, it is essential to break down the signals into their frequency components. This helps to understand the power distribution across different frequency bands, which is crucial to identify the underlying patterns and energy distribution in the brain. Several key features in this analysis, including Power Spectral Density (PSD), Frequency Bands, and Frequency Peaks, offer a detailed understanding of the spectral characteristics of EEG signals. A significant metric in this analysis is the Alpha to Beta Ratio of Band Power, which indicates the balance between the alpha (8-13 Hz) and beta (13-30 Hz) frequency bands in EEG signals. This ratio is insightful for evaluating the equilibrium between relaxed and active mental states. Additionally, Band Power Gamma reflects the power of EEG signals in the gamma frequency band (30-40 Hz) and provides information about cognitive processes, perception, and higher-level brain functions. Moreover, Band Power in specific frequency bands such as Beta, Alpha, Theta, and Delta sheds light on cognitive engagement, relaxation, memory processes, and deep sleep. Together, these features contribute to a comprehensive understanding of the frequency domain and its implications for the analysis of brain activity.

3.3 3.3 Statistical Features

Statistical features in electroencephalography (EEG) are quantitative measures that encompass various aspects of recorded electrical brain activity. These features play a crucial role in characterizing different brain states, detecting abnormalities, and extracting relevant information for both research and clinical applications.

The primary aim of these features is to capture the statistical properties inherent in the EEG signal and provide insights into the underlying dynamics of brain function. Examples of statistical features include auto-correlation, cross-correlation, and higher-order statistics such as bi-spectrum or coherency. These metrics contribute to a nuanced understanding of the relationships and patterns within EEG signals. Additionally, statistical transforms such as the arithmetic mean play a pivotal role. This category further encompasses significant measures like standard deviation, variance, median value, skewness, and kurtosis, providing a comprehensive statistical framework to interpret and analyze EEG data effectively.

Table (1) summarizes how different EEG feature extraction methods contribute to machine learning outcomes, highlighting their strengths and limitations in various contexts.

The most common manifestation of epilepsy is aberrant, excessive brain neuronal activity, which frequently causes convulsions. The identification of spikes, sharp waves, and spike-and-wave complexes are important EEG characteristics. It is important to examine the amplitude, frequency, and duration of these spikes. Spectral power in the alpha (8–13 Hz) and beta (13–30 Hz) bands can diminish during seizures, but it usually increases in the theta (4–8 Hz) and delta (0.5–4 Hz) bands in the frequency domain. Aside from approximation entropy and Lyapunov exponents, nonlinear dynamics can also show variations in EEG signal complexity, which are frequently predictive of epileptic activity.

The neurodegenerative illness known as Alzheimer's disease frequently manifests as a broad slowing of EEG signals. The frequency-domain aspects are very significant. Additionally important are connectivity metrics like decreased coherence in the alpha and beta bands between different brain areas. Furthermore, reduced signal complexity is revealed by entropy and complexity metrics like Shannon entropy and Lempel-Ziv complexity, which is indicative of the cognitive decline linked to Alzheimer's disease.

Figure (1) provides the illustration of normal EEG patterns typically observed in a healthy person, showing a balanced distribution of brain waves vs. EEG patterns typically observed in Alzheimer's disease, highlighting the characteristic slowing of signals with dominant theta and delta waves.

The primary impact of Parkinson's disease is on motor function, which is reflected in the EEG with amplitude and frequency variations

Table 1. Different EEG feature extraction methods

Feature Extraction Method	Contribution	Strengths	Limitations
Time Domain Features	Statistical Measures	Computationally efficient, suitable for real-time applications and easy to implement	Limited in capturing complex temporal dynamics and frequency-specific information
Frequency Domain Features	Power Spectral density using Fourier or wavelet transforms	Effective for tasks requiring frequency information, insights into rhythmic brain activities	May overlook non-stationary aspects and computationally intensive
Time-Frequency Domain Features	Combines time and frequency information	Captures transient events and oscillatory patterns. It's very useful for non-stationary signals	Requires careful parameter tuning and computationally demanding
Spatial Features	Spatial distribution analysis	Useful in BCI and motor-intensive tasks	Complex processing, Sensitive to electrode placement and requires many channels
Connectivity Features	Assesses interaction between brain regions using coherence and phase-locking value	Suitable for cognitive studies and disorders and valuable for understanding functional connectivity	High dimensionality, potential for spurious connections and requires careful interpretation
Non-Linear features	Captures complexity and irregularity using entropy measures and fractal dimensions	Useful in distinguishing normal vs. abnormal states and provides insights into chaotic brain activity	Sensitive to noise and computationally complex

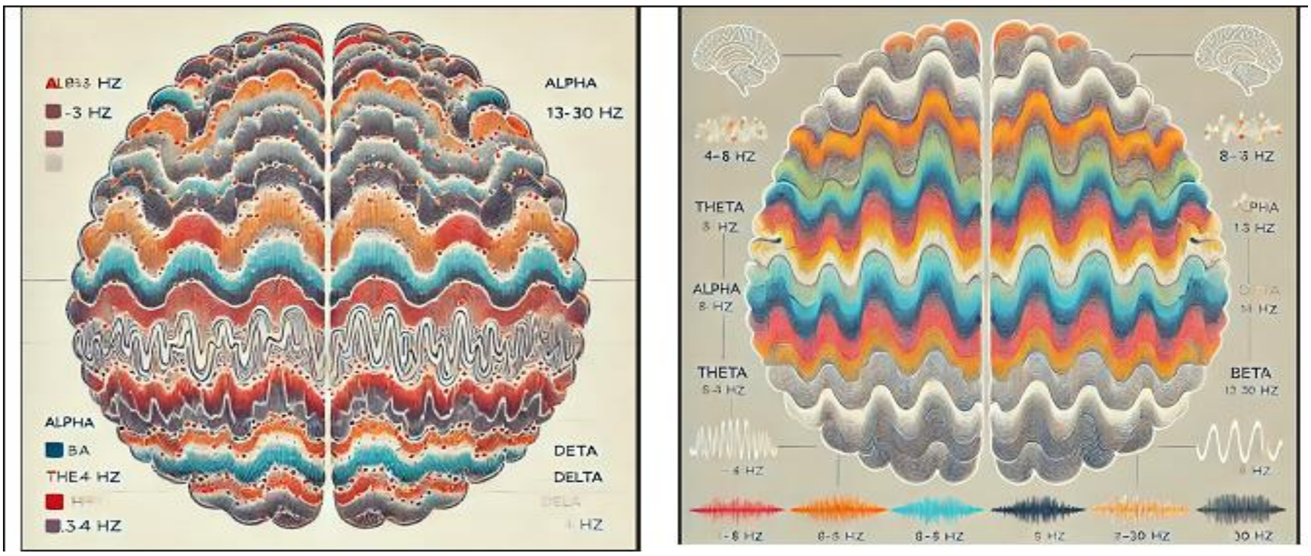


Figure 1. Healthy Control vs. Alzheimer Patients

Table 2. Comparative Analysis of Database Characteristics

Database Name	**Size (No. of Records)**	**Type of Data**	**Sampling Rate**	**Key Features**	**Strengths**	**Limitations**
Database A	5,000	EEG Signals	256 Hz	Raw EEG, Event Markers	High data quality, Extensive metadata	Limited to specific demographics
Database B	10,000	EEG, EOG	512 Hz	Pre-processed Signals	Large sample size, Diverse data types	Data access restrictions
Database C	3,500	EEG, ECG	128 Hz	Annotated Seizures	Contains rare seizure data, Easy access	Lower sampling rate, Smaller size
Database D	7,200	EEG	256 Hz	Artifact-free Data	Clean data, Good for general EEG studies	Lacks physiological diversity

Table 3. Key Comparative Insights across Databases

Insight	**Database A**	**Database B**	**Database C**	**Database D**
Diversity of Data Types	Low	High	Medium	Low
Size vs. Quality Trade-off	Balanced	Large but restricted	Small but specialized	High quality but specific
Ease of Access	Moderate	Restricted	High	Moderate
Sampling Rate Adequacy for EEG Analysis	Adequate	High	Low	Adequate
Demographic Representation	Limited	Broad	Limited	Moderate

Modified brain connection can be revealed by characteristics like phase-amplitude coupling and wavelet coherence. Furthermore, during motor tasks, motor-related potentials such as Event-Related De-Synchronization (ERD) in the beta band may be diminished, showing that the disease impacts motor control.

Table 2 has been revised to include columns for "Strengths" and "Limitations," which provide a quick summary of the advantages and drawbacks of each database. This makes the table more informative and user-friendly, enabling readers to compare the databases at a glance.

Table 3 summarizes key comparative insights across the databases, such as diversity, data quality, and accessibility. This table helps in providing a broader understanding of how the databases compare with one another on important aspects. These tables should help readers better understand the characteristics of the databases and how they compare in terms of various important factors.

4. Implementation

This section explores the way of handling EEG information in detail with respect to preprocessing and various feature extractions.

4.1 EEG Preprocessing

Preparing EEG data for machine learning entails a series of procedures to clean, filter, and prepare the data. These actions are crucial for guaranteeing the accuracy and applicability of the data, which has a direct impact on how well machine learning algorithms operate.

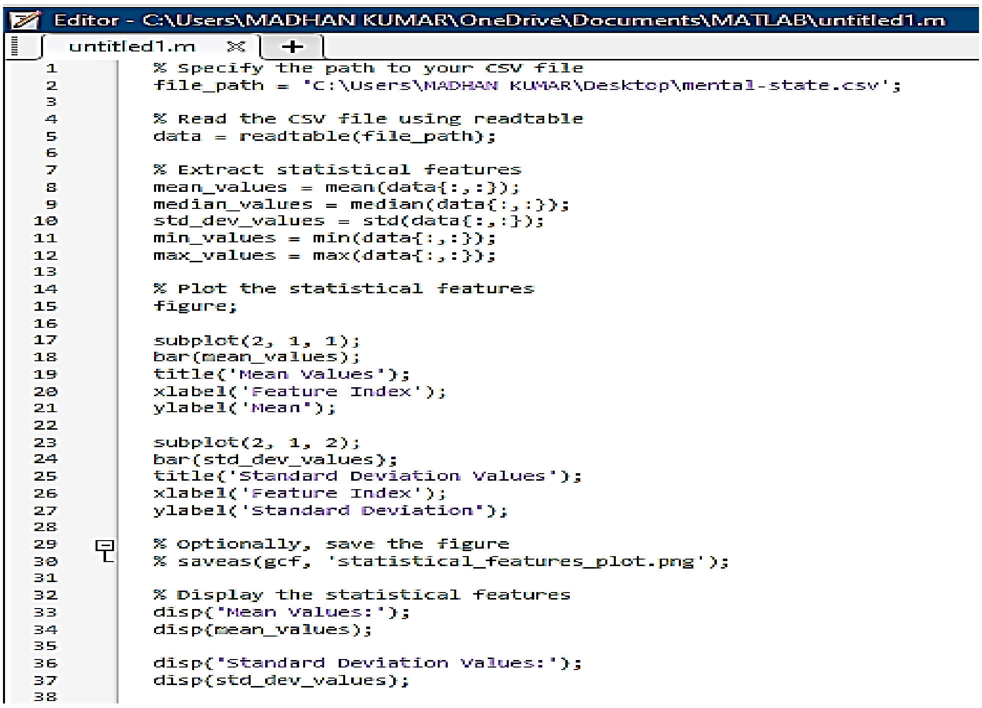
Consider the kind of task, data complexity, labeled data availability, interpretability, computational resources, generalization ability, and training and inference speed when choosing a machine learning algorithm for EEG analysis. To satisfy the particular needs of the application, the best option necessitates striking a balance between these factors.

In order to ensure that raw EEG data is clean and appropriate for machine learning, preprocessing is an essential step in getting the data ready for analysis. Data acquisition is the first step in the procedure, where electrodes are applied to the scalp to record EEG signals. The next step is filtering, where undesirable frequencies, including baseline drift or muscular artifacts, are eliminated by applying band-pass, low-pass, or high-pass filters. Next, artifact removal—which frequently involves the use of techniques like Independent Component Analysis (ICA) or regression-based methods—is carried out to get rid of non-neural noise brought on by eye blinks, muscle movements, or electrical interference. Re-referencing the data standardizes the signals and lowers noise; this is usually accomplished by removing the average signal of a reference electrode.

4.2 Implementation of Feature Extraction Methods

This implementation underscores the significance of employing statistical features in the evaluation of EEG signals, particularly concerning diverse mental states. The methodology integrates a strategic combination of time windowing techniques and meticulous feature selection, proving to be highly effective in categorizing mental states such as relaxation, neutrality, and concentration. By applying both individual and ensemble classification methods to data collected from four specific points on the scalp, this approach transforms the raw information into a nuanced emotional representation, offering valuable insights into the participant's emotional experiences at different junctures. The dataset utilized in this study comprises 87 columns and 217 rows. The first column serves as an identifier for patients with mental health disorders, while the subsequent columns (F1 to F84) likely correspond to specific categories or features within the dataset. The last column plays a crucial role in classifying each data point into distinct mental states, including but not limited to relaxed, focused, or drowsy. This detailed classification enables a granular understanding of the emotional and mental states of the participants, contributing to a more comprehensive assessment of their psychological well-being. Moreover, the systematic analysis of mental states through statistical features provides valuable insights that can enhance human-computer interactions and contribute significantly to the evaluation and improvement of mental health systems.

Figure 2 is positioned above, illustrating the intricate landscape of EEG signal components. The methodology employed involves the calculation of statistical features across diverse frequency bands and time intervals, allowing for a comprehensive understanding of different aspects of EEG signals. This approach holds particular significance in activities related to EEG analysis, such as applications in brain-computer interfaces, sleep staging, seizure detection, and cognitive state assessment. The dataset encompasses additional information, incorporating mean, median, mode, and standard deviation data sets. These statistical measures are instrumental in capturing the nuances and patterns embedded within the EEG signals. Beyond mere representation, these features serve as crucial input parameters for a spectrum of machine learning, pattern recognition, and classification algorithms. This integration of statistical features enhances the robustness and applicability of the EEG analysis, paving the way for advancements in diverse fields leveraging EEG data for a multitude of applications.


Editor - C:\Users\MADHAN KUMAR\OneDrive\Documents\MATLAB\untitled1.m

untitled1.m

```
% Specify the path to your CSV file
file_path = 'C:\Users\MADHAN KUMAR\Desktop\mental-state.csv';

% Read the CSV file using readtable
data = readtable(file_path);

% Extract statistical features
mean_values = mean(data{:,:});
median_values = median(data{:,:});
std_dev_values = std(data{:,:});
min_values = min(data{:,:});
max_values = max(data{:,:});

% Plot the statistical features
figure;

subplot(2, 1, 1);
bar(mean_values);
title('Mean Values');
xlabel('Feature Index');
ylabel('Mean');

subplot(2, 1, 2);
bar(std_dev_values);
title('Standard Deviation Values');
xlabel('Feature Index');
ylabel('Standard Deviation');

% Optionally, save the figure
% saveas(gcf, 'statistical_features_plot.png');

% Display the statistical features
disp('Mean Values:');
disp(mean_values);

disp('Standard Deviation Values:');
disp(std_dev_values);
```

Figure 2. Statistical Methods

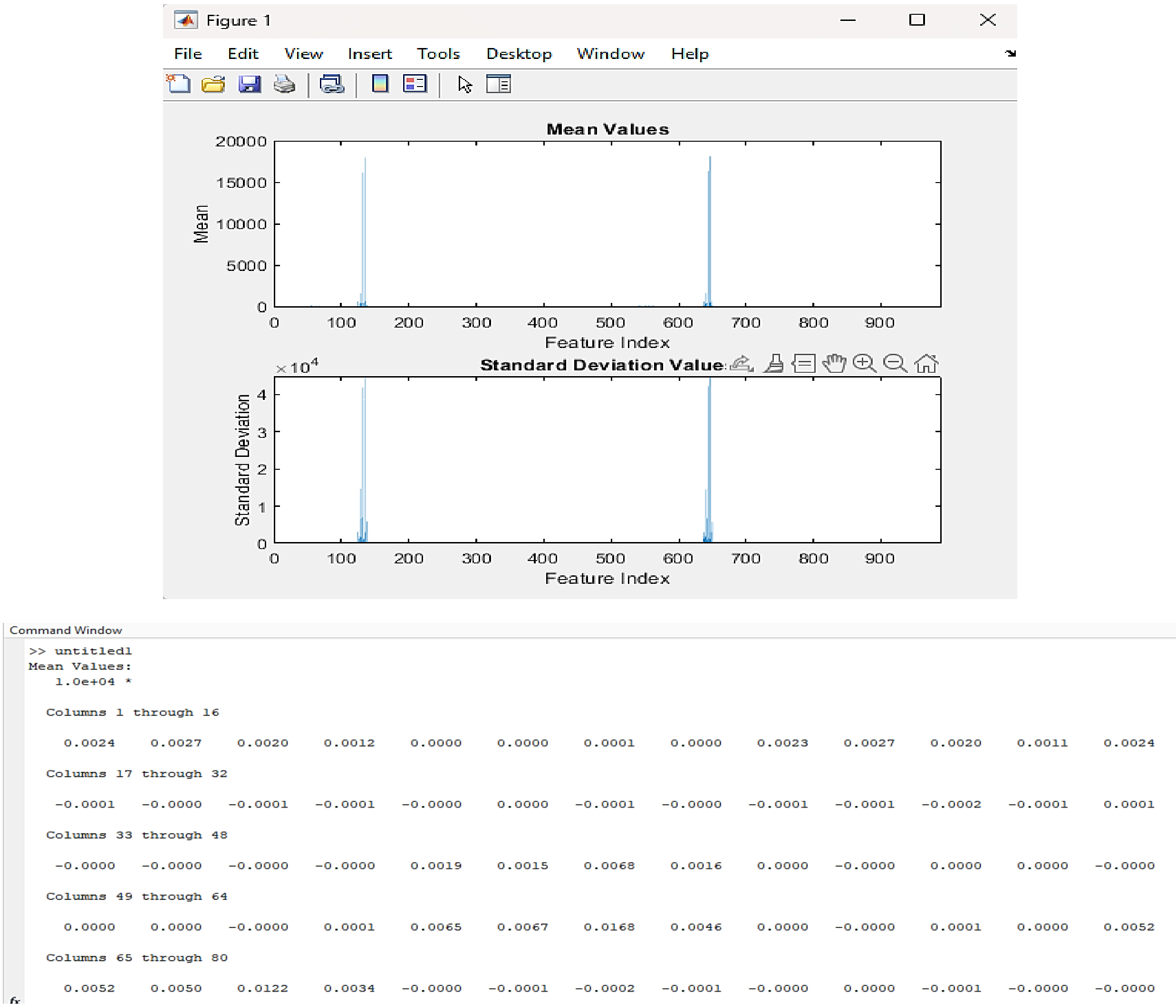


Figure 3. Statistical Features

In Figure 3, the output illustrates precise decimal values derived from the dataset focusing on mental state disorders. This depiction is a result of the utilization of statistical features, showcasing key information such as feature index, mean, and standard deviation. The plotting image provides a visual representation of these essential statistical aspects, contributing to a clearer understanding of the dataset's characteristics related to mental state disorders. The explicit presentation of these features enhances the interpretability and utility of the dataset for further analysis and insights into mental health conditions.

5. Neurological Disease

5.1 Alzheimer's Disease (AD)

EEG-based studies of Alzheimer's disease consistently show decreased alpha power and increased delta and theta activity, indicating cognitive decline. EEG frequency analysis provides valuable biomarkers for diagnosing AD, particularly during early-stage disease progression [18]. Additionally, several studies have emphasized the role of connectivity measures in understanding AD, with reduced synchronization between brain regions serving as a critical diagnostic indicator [19]. In EEG investigations of Alzheimer's patients, several consistent patterns emerge.

5.1.1 Decreased Alpha Wave Activity

Alpha waves, typically found in the 8 to 13 Hz frequency range, tend to decrease in individuals with Alzheimer's disease. These waves are associated with a relaxed wakeful state and are commonly detected when a person has their eyes closed. The diminished alpha wave activity may indicate significant issues in cortical networks, reflecting cognitive processing challenges in Alzheimer's patients.

5.1.2 Increased Theta Wave Activity

Theta waves, ranging between 4 to 7 Hz, exhibit heightened activity in individuals with Alzheimer's disease. These waves are linked to daydreaming, drowsiness, and the early stages of sleep. The increased theta wave activity suggests a state of heightened drowsiness or cognitive impairment in Alzheimer's patients.

5.1.3 Delta Wave Changes

Delta waves, slow brain waves occurring in the 0.5 to 4 Hz frequency range, may show increased activity in individuals with Alzheimer's disease, particularly during deeper sleep stages. The alterations in delta wave activity may vary based on specific brain regions, influencing the disease's progression. These changes in brain wave frequency serve as distinctive markers in Alzheimer's patients compared to healthy individuals, highlighting universal patterns observed in the investigation of Alzheimer's disease. Notably, alpha wave activity declines and increased theta wave activity are characteristic findings, shedding light on potential disruptions in cortical networks and cognitive processing in Alzheimer's patients. Additionally, changes in delta wave activity during deep sleep stages further contribute to our understanding of the disease's impact on brain waves.

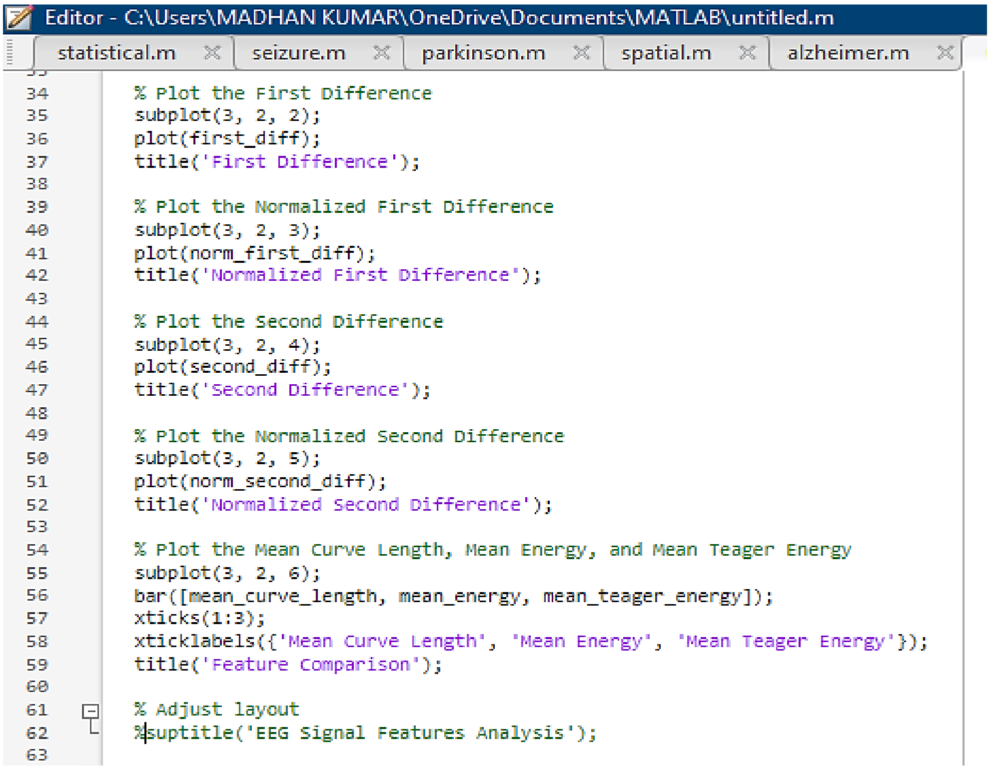


Figure 4. Code for Alzheimer's disease Feature Extraction

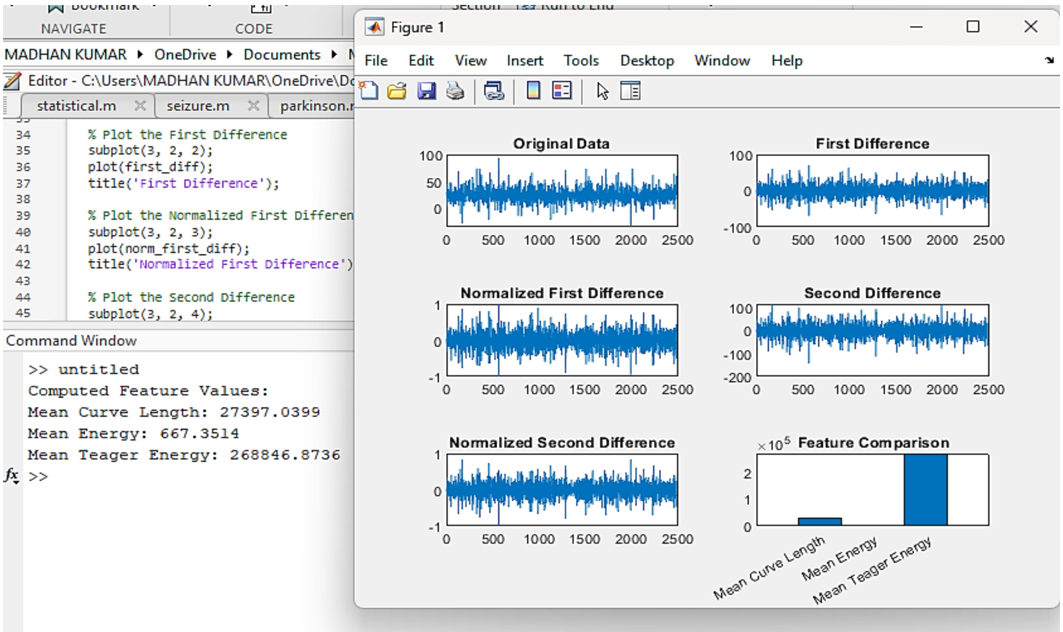


Figure 5. Time Domain Features

Figure 4 provides a comprehensive overview of the Alzheimer's disease dataset, illustrating the application of feature analysis and null value processing to represent the data effectively. Through meticulous charting techniques, the figure captures the intricate details of the dataset's column names, showcasing the richness of information encapsulated within. This visualization aids in understanding the dataset's structure, allowing for informed feature analysis and ensuring that null values are appropriately handled. The detailed extraction of column names enhances the interpretability of the dataset, facilitating further exploration and analysis.

The Alzheimer's illness dataset is used in Figure 5 to visualize age and other disintermediation variables. The mean and frequency are displayed by bar charting.

5.2. Parkinson's Disease (PD)

Parkinson's disease, a neurodegenerative condition primarily affecting the motoric system and associated with dopaminergic cell dysfunction in the substantia nigra region of the brain, is the focal point of this study. Unlike certain neurological conditions, Parkinson's disease is not commonly linked with specific brain wave patterns. However, the investigation reveals distinctive EEG patterns associated with the disease, providing valuable insights into the neurophysiological changes induced by Parkinson's. One notable characteristic of Parkinson's disease is the heightened beta wave oscillations, typically ranging between 13 to 30 Hz, particularly in specific brain regions like the basal ganglia. This increased beta wave activity in the basal ganglia is closely linked to the modulation of motor symptoms such as tremors, stiffness, and bradykinesia. The aberrant beta wave activity may significantly impact the motor functions of individuals with Parkinson's, affecting their ability to control movements.

In addition to alterations in beta wave activity, Parkinson's disease manifests significant changes in other frequency bands, including alpha and theta waves.

```
statistical.m  seizure.m  parkinson.m  spatial.m  alzheimer.m

% Load EEG dataset
dataset = readmatrix("C:\Users\MADHAN KUMAR\Downloads\parkinsons.csv");

% Assume the dataset represents EEG data sampled at a certain frequency (Fs)
Fs = 1000; % Replace with your actual sampling frequency

% Check for NaN values and replace with zeros
dataset(isnan(dataset)) = 0;

% Define frequency bands
alpha_band = [8 13]; % in Hz
beta_band = [13 30]; % in Hz
gamma_band = [30 40]; % in Hz
delta_band = [0.5 4]; % in Hz
theta_band = [4 8]; % in Hz

% Calculate power in each frequency band
alpha_power = bandpower(dataset, Fs, alpha_band);
beta_power = bandpower(dataset, Fs, beta_band);
gamma_power = bandpower(dataset, Fs, gamma_band);
delta_power = bandpower(dataset, Fs, delta_band);
theta_power = bandpower(dataset, Fs, theta_band);

% Display the calculated powers
disp(['Alpha Power: ', num2str(alpha_power)]);
disp(['Beta Power: ', num2str(beta_power)]);
disp(['Gamma Power: ', num2str(gamma_power)]);
disp(['Delta Power: ', num2str(delta_power)]);
disp(['Theta Power: ', num2str(theta_power)]);

% Additional statistical features
mean_val = mean(dataset);
variance = var(dataset);
std_deviation = std(dataset);
skewness_val = skewness(dataset);
kurtosis_val = kurtosis(dataset);
rms = sqrt(mean(dataset.^2));

% Print the calculated features
disp(['Mean: ', num2str(mean_val)]);
disp(['Variance: ', num2str(variance)]);
```

Figure 6. Code for Parkinson Disease Feature Extraction

```
>> parkinson
Alpha Power: 0      6146.21975      11852.9029      3055.90198  1.57634391e-05  8.23992743e-10  5.33870685e-06  4.402
Beta Power: 0      256.19273      1709.7357      288.97992  4.1090285e-06  1.9727747e-10  1.7620418e-06  1.1944683e-0
Gamma Power: 0      533.97095      2120.0783      115.63881  4.6659322e-06  3.1250994e-10  1.8514533e-06   1.233809e-
Delta Power: 0      23453.3219      40020.7068      12524.9977  4.58813051e-05  2.29663896e-09  1.38764698e-05  1.317
Theta Power: 0      23510.1857      40339.9241      12578.9737  4.61946727e-05  2.33553784e-09  1.40704209e-05  1.325
Mean: 0      154.2286      197.1049      116.3246    0.006220462  4.395897e-05    0.00330641    0.003446359    0.0099199
Variance: 0      1713.1375      8370.7033      1894.1134    2.35044e-05  1.2125653e-09    8.807685e-06  7.6119521e-06
Standard Deviation: 0      41.3901      91.4915      43.5214    0.00484813  3.48219e-05    0.00296777    0.00275898    0.
Skewness: NaN      0.5871759      2.522549      1.207966      3.061165      2.62865      3.334801      3.050196
Kurtosis: NaN      2.357474      10.40246      3.607337      14.69405      13.56173      16.82134      14.62874
Root Mean Square: 0      159.6585      217.2053      124.1605    0.00787896  5.602445e-05    0.00443789    0.00441025
>>
```

Figure 7. Frequency and Statistical Feature

These modifications occur in brain regions associated with attention, cognition, and sensory processing. Individuals with Parkinson's often exhibit reduced alpha wave activity (8-13 Hz) and irregular theta wave patterns (4-7 Hz) in these specific regions. These alterations are thought to play a pivotal role in the development of non-motor symptoms such as cognitive impairments and sensory abnormalities commonly observed in Parkinson's disease. Parkinson’s disease is primarily characterized by motor dysfunction, but EEG studies have uncovered significant neurophysiological changes. PD patients exhibit altered beta and gamma rhythms, which are linked to both motor and non-motor symptoms [20]. In particular, beta oscillations are associated with motor control deficits, while gamma activity correlates with cognitive impairments [21].

Figure 6 illustrates the implementation of the Parkinson's disease dataset, showcasing the code associated with common wave terms. The plot displays estimated variance, kurtosis, skewness, and Root Mean Square (RMS) values, providing a comprehensive visualization of key features. The code extraction was accomplished through recognition algorithms, enhancing the accessibility and understanding of the underlying dataset and its associated code. This representation aids researchers and practitioners in gaining insights into the essential characteristics of the Parkinson's disease dataset and facilitates further analysis and interpretation of the provided information.

Figure 7 provides a visual representation of brain wave patterns extracted from datasets related to Parkinson's disease. The illustration highlights key statistical metrics, including mean, variance, skewness, and kurtosis. These metrics offer valuable insights into the distinctive characteristics of brain wave signals associated with Parkinson's disease. The comprehensive analysis depicted in the figure contributes to a better understanding of the neurophysiological aspects of Parkinson's disease, aiding researchers and clinicians in identifying relevant patterns and potential markers associated with the condition.

5.3 Seizure Disorder

Seizures can lead to diverse abnormal brain wave patterns, influenced by seizure type and specific affected brain regions. Focal seizures may show localized disruptions, while generalized seizures impact both hemispheres simultaneously, resulting in widespread abnormal brain wave activity. The unique patterns observed during seizures arise from the combination of seizure type and affected brain areas. During generalized tonic-clonic seizures involving both brain sides, abnormal patterns include a combination of high-amplitude, fast activity (beta waves) and low-amplitude, slow activity (delta waves). The simultaneous presence of beta and delta waves indicates widespread electrical discharges across the entire brain during seizures. Absence seizures feature a characteristic pattern called generalized 3-per-second spike and wave, indicative of synchronous aberrant neural activity and momentary loss of awareness. EEG remains the gold standard for diagnosing and monitoring seizure disorders. EEG patterns during seizures, such as high-amplitude sharp waves, are key features for identifying epilepsy types [22]. Further, deep learning models applied to EEG data have demonstrated high accuracy in predicting seizures before they occur, providing a significant advancement in epilepsy management [23]. Focal seizures, originating from specific brain regions, exhibit various patterns depending on the involved area. Motor symptoms in focal seizures may present rhythmic activity or spikes within the motor cortex, responsible for movement. Sensory symptoms in focal seizures can be detected by the sensory cortex, disrupting sensory processing during seizures, evident on the EEG. Focal seizures with impaired awareness manifest diverse abnormal brain wave patterns contingent on the specific brain region affected, highlighting complex alterations in neural activity underlying impaired consciousness during these seizures.

Figure 8 portrays a meticulous implementation of seizure disorder detection using advanced techniques like Support Vector Machines (SVMs). The methodology entails systematically labeling a trained dataset, allowing the SVM algorithm to discern patterns and relationships within the data.

```
% Open the EEG dataset
% Replace 'path_to_dataset' with the actual path to your dataset file
dataset = readmatrix("C:\Users\MADHAN KUMAR\Desktop\MCA\KEC PROJECTS\Mini " + ...
    "project sem-2\mini_project\Epileptic Seizure Recognition.csv");

% Assume the dataset represents EEG data sampled at a certain frequency (Fs)
Fs = 1000; % Replace with your actual sampling frequency

% Calculate the power spectral density (PSD) using periodogram
[Pxx, frequencies] = periodogram(dataset, [], length(dataset), Fs);

% Define frequency bands
alpha_band = [8 18]; % in Hz
beta_band = [13 30]; % in Hz
gamma_band = [30 40]; % in Hz
delta_band = [0.5 4]; % in Hz
theta_band = [4 8]; % in Hz

% Find indices corresponding to each frequency band
alpha_indices = find(frequencies >= alpha_band(1) & frequencies <= alpha_band(2));
beta_indices = find(frequencies >= beta_band(1) & frequencies <= beta_band(2));
gamma_indices = find(frequencies >= gamma_band(1) & frequencies <= gamma_band(2));
delta_indices = find(frequencies >= delta_band(1) & frequencies <= delta_band(2));
theta_indices = find(frequencies >= theta_band(1) & frequencies <= theta_band(2));

% Calculate power in each frequency band
alpha_power = sum(Pxx(alpha_indices));
beta_power = sum(Pxx(beta_indices));
gamma_power = sum(Pxx(gamma_indices));
delta_power = sum(Pxx(delta_indices));
theta_power = sum(Pxx(theta_indices));

% Display the calculated powers
disp(['Alpha Power: ', num2str(alpha_power)]);
disp(['Beta Power: ', num2str(beta_power)]);
disp(['Gamma Power: ', num2str(gamma_power)]);
disp(['Delta Power: ', num2str(delta_power)]);
disp(['Theta Power: ', num2str(theta_power)]);

% Continue with your existing code for mean, variance, etc.
mean_val = mean(dataset);
```

Figure 8. Code for Seizure Disease Feature Extraction

```
>> seizure
Alpha Power: 6.5217
Beta Power: 8.5933
Gamma Power: 4.9105
Delta Power: 2.4678
Theta Power: 2.4856
Mean: 23.584351       27.060411       20.4529311       11.5260439      0.0144492414      0.00339712131
Variance: 111.17462331817       588.17640720619       5199.0500516478       368.51980790084       179.
Standard Deviation: 10.5439378       24.2523485       72.1044385       19.1968697       13.382998
Skewness: 0.333754      -1.21146      0.267304      -4.60608      0.778951      -4.37201      0.580319
Kurtosis: 6.7734075       101.97128       19.861421       42.672314       12.322353       149.38195
Root Mean Square: 25.8331452       36.3346251       74.93514       22.3879619       13.3803062
>>
```

Figure 9. Statistical and Frequency Features

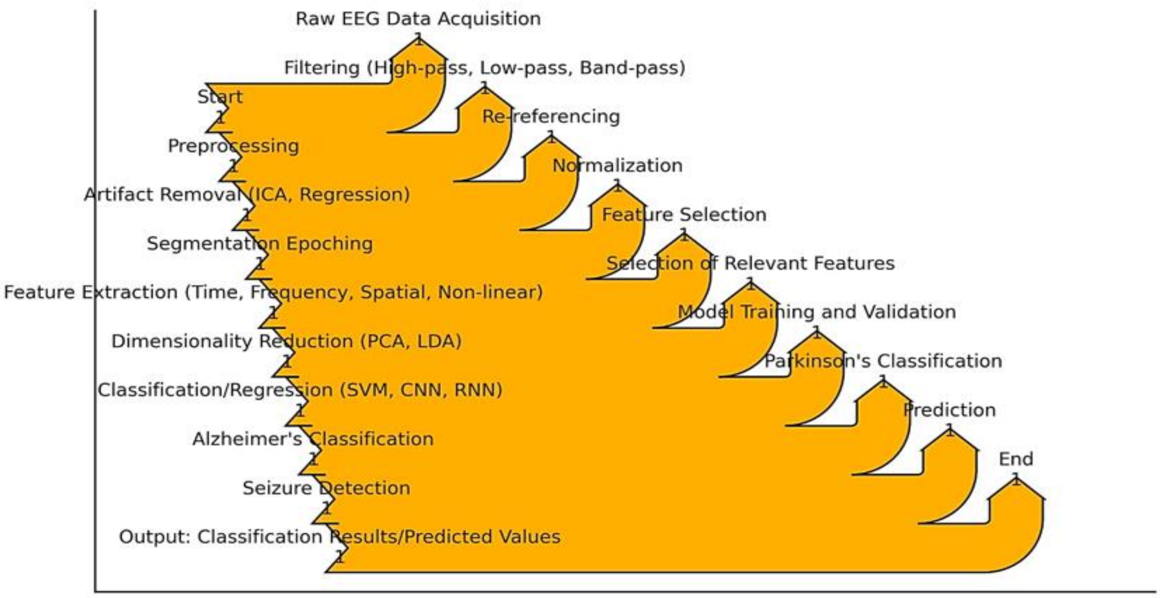


Figure 10. Algorithmic Flow of EEG Classification

This approach enhances the precision of seizure detection by leveraging the capabilities of SVMs to classify intricate patterns associated with abnormal brain wave activity indicative of seizures.

Figure 9 illustrates the calculation of data counts and duplicate values using the seizure disorder dataset and SVM approaches. This depiction provides insights into the quantitative aspects of the dataset, offering a comprehensive view of the data distribution and the utilization of SVM techniques in managing and processing the information.

Figure 10 shows a flow diagram of EEG classification from data acquisition, preprocessing, feature extraction and selection.

6. Conclusion

In the field of neurological research, extracting features from EEG signals is a crucial technique for understanding and diagnosing various brain disorders. EEG signals contain valuable information that provides insights into the complex electrical processes within the brain. This study explores various features, including time domain attributes, statistical metrics, and frequency characteristics, to identify patterns associated with conditions such as seizures, Alzheimer's, and Parkinson's diseases. By carefully analyzing these features using computational tools like MATLAB, healthcare professionals can identify irregularities and make accurate diagnoses. Statistical features, such as mean, variance, skewness, and kurtosis, are particularly effective in distinguishing mental state disorders through the analysis of EEG signals. A comprehensive dataset further supports the effectiveness of these statistical features, demonstrating their importance in enhancing diagnostic capabilities. This investigation establishes a strong foundation using feature extraction and selection methods to improve the prediction and understanding of mental state disorders through the analysis of EEG signals.

7. Future Enhancements

In this research, further improvements can be made by integrating appropriate feature-based techniques in EEG signals with proper preprocessing to enhance the model's predictive capabilities. Advancements in signal processing and computational methodologies offer opportunities for a more in-depth understanding of brain waves. The use of emerging technologies such as machine learning, deep learning and precision healthcare holds significant potential in improving diagnostic frameworks. By including advanced deep learning algorithms, particularly those based on reinforcement learning, intricate patterns in neurological disorders can be revealed in the dynamic real-time environment. Collaborating with fields like neuroinformatics and data science can result in comprehensive diagnostic models. It is important to validate and refine models across diverse demographic groups and larger datasets to improve generalizability and robustness, especially when dealing with data variability. Above all, ensure the ethical use of sensitive medical information in all forthcoming research initiatives.

References

[1] F. Castro, J. Melgarejo, C.A. Chavez, G.A. De Erausquin, J.D. Terwilliger, J.H. Lee, G.E. Maestre, Total plasma homocysteine and depressive symptoms in older Hispanics. Journal of Alzheimer's Disease, 82(s1), (2021) S263-S269. https://doi.org/10.3233/JAD-201062

[2] G. Melikishvili, T. Bienvenu, N. Tabatadze, T. Gachechiladze, E. Kurua, S. Gverdtsiteli, M. Melikishvili, O. Dulac, Novel UBE3A pathogenic variant in a large Georgian family produces non-convulsive status epilepticus responsive to ketogenic diet. Seizure. 94, (2022) 70-73. https://doi.org/10.1016/j.seizure.2021.11.012

[3] B. Gosala, P.D. Kapgate, P. Jain, R. Chaurasia, M. Gupta, Wavelet transforms for feature engineering in EEG data processing: An application on Schizophrenia. Biomedical Signal Processing and Control, 85, (2023) 104811. https://doi.org/10.1016/j.bspc.2023.104811

[4] S.J.J. Jui, R.C. Deo, P.D. Barua, A. Devi, J. Soar, U.R. Acharya, Application of entropy for automated detection of neurological disorders with electroencephalogram signals: a review of the last decade (2012-2022). IEEE Access, (2023) 71905-71924. https://doi.org/10.1109/ACCESS.2023.3294473

[5] A. Montenegro, G. Sosa, N. Figueroa, V. Vargas, H. Franco, Evaluation of stabilometry descriptors for human balance function classification using diagnostic and statokinesigram data. Biomedical Signal Processing and Control, 84, (2023) 104861. https://doi.org/10.1016/j.bspc.2023.104861

[6] J. Li, W. Pan, H. Huang, J. Pan, F. Wang, STGATE: Spatial-temporal graph attention network with a transformer encoder for EEG-based emotion recognition. Frontiers in Human Neuroscience, 17, (2023) 1169949. https://doi.org/10.3389/fnhum.2023.1169949

[7] H. Yu, S. Baek, J. Lee, I. Sohn, B. Hwang, C. Park, Deep Neural Network-based Empirical Mode Decomposition for Motor Imagery EEG Classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering, IEEE, (2024). https://doi.org/10.1109/TNSRE.2024.3432102

[8] S.Y. Shah, H. Larijani, R.M. Gibson, D. Liarokapis, Epileptic seizure classification based

on random neural networks using discrete wavelet transform for electroencephalogram signal decomposition. Applied Sciences, 14(2), (2024) 599.https://doi.org/10.3390/app14020599

[9] R. Koliqi, A. Fathima, A.K. Tripathi, N. Sohi, R.E. Jesudasan, C. Mahapatra, Innovative and Effective Machine Learning-Based Method to Analyze Alcoholic Brain Activity with Nonlinear Dynamics and Electroencephalography Data. SN Computer Science, 5(1), (2023) 113. https://doi.org/10.1007/s42979-023-02424-6

[10] M, Ragavi, M.K. Subha Sri Lakshmi, Automation by Brain Sense through Eeg Waves. International Research Journal of Multidisciplinary Technovation, 2(4), (2020) 1-3. https://doi.org/10.34256/irjmt2041

[11] S.A. Asha, C. Sudalaimani, P. Devanand, G. Alexander, A.M. Lathikakumari, S.V. Thomas, R.N. Menon, Analysis of EEG microstates as biomarkers in neuropsychological processes–Review. Computers in Biology and Medicine, 173, (2024) 108266. https://doi.org/10.1016/j.compbiomed.2024.108266

[12] S. Liu, J. Wang, S. Li, L. Cai, Epileptic seizure detection and prediction in EEGS using power spectra density parameterization. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 31, (2023) 3884 - 3894. https://doi.org/10.1109/TNSRE.2023.3317093

[13] Y. Ezazi, P. Ghaderyan, A new Cepstral-based biomarker of reward positivity evaluated in Parkinson’s disease detection. Tabriz Journal of Electrical Engineering, (2024). https://doi.org/10.22034/tjee.2024.60087.4798

[14] Y. Wang, C. Song, T. Zhang, Z. Yao, Z. Chang, D. Wang, Feature Extraction of Motor Imagery EEG via Discrete Wavelet Transform and Generalized Maximum Fuzzy Membership Difference Entropy: A Comparative Study. Electronics, 12(10), (2023) 2207. https://doi.org/10.3390/electronics12102207

[15] A. Maturana-Candelas, C. Gómez, J. Poza,S.J. Ruiz-Gómez, R. Hornero, Inter-band Bispectral Analysis of EEG Background Activity to Characterize Alzheimer's Disease Continuum. Frontiers in Computational Neuroscience, 14, (2020) 70. https://doi.org/10.3389/fncom.2020.00070

[16] B.K. Gulay, N. Demirel, A. Vahaplar, C. Guducu, A novel feature extraction method using chemosensory EEG for Parkinson's disease classification. Biomedical Signal Processing and Control, 79(2), (2023) 104147. https://doi.org/10.1016/j.bspc.2022.104147

[17] F.R. Farina, D.D. Emek-Savaş, L. Rueda-Delgado, R. Boyle, H. Kiiski, G. Yener, R. Whelan, A comparison of resting state EEG and structural MRI for classifying Alzheimer's disease and mild cognitive impairment. Neuroimage. 215 (2020) 116795. https://doi.org/10.1016/j.neuroimage.2020.116795

[18] L.V. Tran, H.M. Tran, T.M. Le, T.T.M. Huynh, H.T. Tran, S.V.T. Dao, Application of Machine Learning in Epileptic Seizure Detection. Diagnostics (Basel), 12(11), (2022) 2879. https://doi.org/10.3390/diagnostics12112879

[19] R. Naily, S. Yahia, M. Zaied, A New Deep Learning Architecture Based on LSTM and Wavelet Transform for Epileptic EEG Signal Classification. International Conference on Intelligent Systems Design and Applications, (2023), 353-362. https://doi.org/10.1007/978-3-031-64813-7_36

[20] I. Ahmad, X. Wang, D. Javeed, P. Kumar, O.W. Samuel, S. Chen, (2023) A hybrid deep learning approach for epileptic seizure detection in EEG signals. IEEE Journal of Biomedical and Health Informatics, IEEE, 1-12. https://doi.org/10.1109/JBHI.2023.3265983

[21] K. Baik, J.H. Jung, S.H. Jeong, S.J. Chung, H.S. Yoo, P.H. Lee, Y.H. Sohn, S.W. Kang, B.S. Ye, Implication of EEG theta/alpha and theta/beta ratio in Alzheimer's and Lewy body disease. scientific reports, 12(1), (2022) 18706. https://doi.org/10.1038/s41598-022-21951-5

[22] M. Ljubicic, S. Sare, I. Kolcic, Sleep Quality and Evening Salivary Cortisol Levels in Association with the Psychological Resources of Parents of Children with Developmental Disorders and Type 1 Diabetes. Journal of autism and developmental disorders, (2024) 1-4. https://doi.org/10.1007/s10803-024-06269-7

[23] M. Zhu, A. HajiHosseini, T.R. Baumeister, S. Garg, S. Appel-Cresswell, M.J. McKeown, Altered EEG alpha and theta oscillations characterize apathy in Parkinson's disease during incentivized movement. NeuroImage: Clinical, 23, (2019) 101922. https://doi.org/10.1016/j.nicl.2019.101922

Authors Contribution Statement

K. Nanthini: Conceptualization, Formal analysis, Methodology and Implementation, Writing – original draft, Writing – review and editing; D. Sivabalaselvamani: Supervision, Methodology, Writing Draft, Review and Editing; M.C. Madhan Kumar: Writing Draft; R. Kaviya: Writing Draft, All the authors read and approved the final version of the manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article. To the best of my knowledge and belief, any actual, perceived, or potential conflicts between my duties as an employee and my private and/or business interests have been fully disclosed in this form by the requirements of the journal.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License