Classification of EEG Signals Utilizing DWT for Feature Extraction and Evolutionary Algorithms for Feature Selection

Mayur Akewar ¹

¹Shri Ramdeobaba College of Engineering and Management

December 7, 2023

Abstract

This paper introduces an EEG signal classification approach, leveraging machine learning algorithms. The methodology involves the extraction of features from EEG signal datasets through discrete wavelet transform (DWT). Optimal feature selection is then accomplished using evolutionary algorithms, specifically Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). To identify the most effective classification method, various machine learning algorithms, including Support Vector Machines (SVM), Naive Bayes, Decision Trees, and Random Forest, are systematically compared. This comprehensive evaluation aims to enhance the accuracy and efficiency of EEG signal classification for improved diagnosis and understanding of neurological conditions.
Classification of EEG Signals Utilizing DWT for Feature Extraction and Evolutionary Algorithms for Feature Selection

Mayur Akewar
Alumni, Department of Computer Science & Engineering
Shri Ramdeobaba College of Engineering & Management
Email: mayurakewar87@gmail.com

Abstract—The electroencephalogram (EEG) signals, which capture the electrical activities of the brain, serve as a valuable tool for understanding neuronal function. The analysis of EEG signals is pivotal for detecting and diagnosing various brain disorders, with a specific focus on conditions such as epilepsy and seizures. This paper introduces an EEG signal classification approach, leveraging machine learning algorithms. The methodology involves the extraction of features from EEG signal datasets through discrete wavelet transform (DWT). Optimal feature selection is then accomplished using evolutionary algorithms, specifically Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). To identify the most effective classification method, various machine learning algorithms, including Support Vector Machines (SVM), Naive Bayes, Decision Trees, and Random Forest, are systematically compared. This comprehensive evaluation aims to enhance the accuracy and efficiency of EEG signal classification for improved diagnosis and understanding of neurological conditions. The code is published at the URL [https://github.com/akewarmayur/EEGSignalClassification]

Index Terms—EEG, Discrete Wavelet Transform, Genetic Algorithm, Particle Swarm Optimization, Machine Learning.

I. INTRODUCTION

The human brain’s intricate electrical activities are encapsulated through electroencephalogram (EEG) signals, providing an invaluable gateway into understanding neuronal function. This paper delves into the realm of EEG signal analysis, an essential endeavor for the detection and diagnosis of various brain disorders, with a specific emphasis on conditions such as epilepsy and seizures. The complexity and richness of EEG signals necessitate advanced analytical tools, prompting the introduction of a novel EEG signal classification approach that harnesses the power of machine learning algorithms. EEG signals, as snapshots of brain activity, present a wealth of information crucial for unraveling the mysteries of neurological conditions. The benefits of EEG signal classification are profound, serving as a diagnostic beacon that illuminates the path toward accurate identification and understanding of disorders affecting the brain. The integration of machine learning (ML) and artificial intelligence (AI) in EEG signal classification emerges as a pivotal advancement. These technologies bring sophistication to the analysis, enabling a more nuanced interpretation of intricate EEG patterns.

The methodology adopted in this paper intricately involves discrete wavelet transform (DWT) for extracting features from EEG signal datasets. DWT provides a robust framework for unveiling essential patterns within the signals, contributing to a more comprehensive analysis. Evolutionary algorithms, notably Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), take center stage for optimal feature selection. This strategic approach ensures that the features selected for classification enhance the model’s discriminative power. The arsenal of machine learning algorithms deployed in this study encompasses Support Vector Machines (SVM), Naive Bayes, Decision Trees, and Random Forest. These algorithms are systematically compared and evaluated to identify the most effective classification method. This rigorous analysis seeks not only to enhance the accuracy and efficiency of EEG signal classification but also to foster a deeper understanding of neurological conditions. The journey outlined in this paper represents a significant stride towards advancing the field, with implications for more precise diagnoses and improved insights into the intricate landscape of brain disorders.

This paper represents a significant stride in exploring EEG signals, merging sophisticated analytical techniques, machine learning, and evolutionary algorithms. The ultimate goal is to not only enhance the accuracy of EEG signal classification but also to deepen our understanding of neurological conditions, paving the way for more effective diagnoses and targeted interventions.

II. LITERATURE REVIEW

In the vast landscape of neuroscience, the classification of EEG signals stands as a crucial frontier, with machine learning methods emerging as powerful allies in unraveling the intricate patterns within these brainwave recordings. This literature review delves into the diverse methodologies and advancements employed in EEG signal classification. In [1], the authors address the critical problem of EEG signal analysis for detecting epileptic seizures in health informatics. Traditionally, EEG signal classification involves two key steps: signal preprocessing and classification. While preprocessing aims to reduce noise and signal dimensionality, it inevitably
leads to a loss of information before classification. In this study, the authors propose an innovative approach considering the loss of information in the classification step. The novelty lies in incorporating a fuzzy classifier to account for the uncertainty caused by the loss of information during dimensionality reduction. To integrate the fuzzy classifier, the preprocessing step is modified to include a fuzzification procedure. The Fuzzy Decision Tree (FDT) is employed as the fuzzy classifier for epileptic seizure detection, achieving an impressive 99.5% classification accuracy. Comparative analysis with other studies underscores the effectiveness of FDT in detecting epileptic seizures, demonstrating its potential as a valuable tool in EEG signal classification. The current focus in brain-computer interface (BCI) research is on motor imagery (MI) electroencephalogram (EEG) signal classification, with deep learning emerging as a promising method for automatic feature extraction and classification. However, challenges arise in practical applications due to the need for a large amount of labeled data and the computational expense of training deep learning models from scratch. To address these issues, the paper [2] introduces a deep transfer convolutional neural network (CNN) framework based on VGG-16 for EEG signal classification. This framework utilizes a pre-trained VGG-16 CNN model on ImageNet as a base and transfers its parameters to a target CNN model for MI EEG signal classification. The target model’s front-layer parameters are frozen, while later-layer parameters are fine-tuned using the target MI dataset. The proposed framework, validated on the BCI competition IV benchmark dataset 2b, demonstrates improved accuracy and efficiency in EEG signal classification compared to traditional methods like support vector machine (SVM), artificial neural network (ANN), and standard CNN. In [3], the authors present a novel approach, termed epileptic EEG signal classification (EESC), for accurate categorization of different epileptic states using electroencephalogram (EEG) signals. The methodology involves transforming epileptic EEG signals into power spectrum density energy diagrams (PSDEDs), followed by the application of deep convolutional neural networks (DCNNs) and transfer learning to automatically extract features from these PSDEDs. The proposed approach exhibits superior performance in comparison to existing methods, demonstrating high accuracy and efficiency. In a case study using CHB-MIT epileptic EEG data, the EESC method achieves an average classification accuracy exceeding 90%, highlighting its effectiveness in epilepsy analysis. In [4], the authors explore the application of neural networks (NN) in EEG signal classification, comparing the performance of SVM, logistic regression, and NN machine learning algorithms. They introduce two-layer LSTM and four-layer improved NN deep learning algorithms, featuring novel one-dimensional gradient descent activation functions with radial basis operations in the initial layers for enhanced performance. Statistical features such as mean, standard deviation, kurtosis, and skewness are extracted from EEG data sourced from the Bonn database and applied to different classification techniques. Evaluation metrics include accuracy, precision, recall, and F1 score. The improved NN and LSTM architectures outperform others, with extensive simulations involving various activation functions, optimizers, and loss models conducted using Python in Keras. In [5], the authors introduce a privacy-preserving deep learning (DL) architecture called Federated Transfer Learning (FTL) for electroencephalographic (EEG) classification in Brain-Computer Interfaces (BCI). Due to privacy concerns, constructing large EEG-BCI datasets is challenging. FTL operates within the federated learning framework, utilizing the single-trial covariance matrix to extract discriminative information from multi-subject EEG data through domain adaptation techniques. Evaluating the architecture on the PhysioNet dataset for 2-class motor imagery classification, FTL achieves a 2% higher classification accuracy in subject-adaptive analysis without actual data sharing. In the absence of multi-subject data, FTL outperforms other state-of-the-art DL architectures by 6% in accuracy. In [6], the authors tackle the challenging task of epileptic seizure detection by classifying electroencephalography (EEG) signals into ictal and interictal classes. Given the complex and non-linear nature of EEG signals, extracting discriminative features is crucial for accurate classification. This paper introduces an approach that combines fuzzy-based and traditional machine learning algorithms to address this issue. The proposed framework successfully classifies unknown EEG signal segments into ictal and interictal classes, with empirical evaluation on benchmark datasets (Bonn and CHB-MIT) showing that K-Nearest Neighbor (KNN) and Fuzzy Rough Nearest Neighbor (FRNN) achieve the highest classification accuracy, along with improved sensitivity and specificity percentages. In [7], the authors propose a multilevel machine learning approach for diagnosing epilepsy using electroencephalography (EEG) signals. The method comprises pre-processing, feature extraction, feature concatenation, feature selection, and classification phases. Tunable-Q wavelet transform (TQWT) is employed for pre-processing, generating 25 frequency coefficient sub-bands. Feature extraction utilizes quadruple symmetric pattern (QSP), extracting 256 features from the raw EEG signal and the 25 sub-bands. Neighborhood component analysis (NCA) is applied for feature selection, choosing the most significant features. The classification phase employs the k nearest neighbors (kNN) classifier. Tested on the Bonn EEG dataset, the proposed method achieves a 98.4% success rate for a case with five classes, demonstrating its potential for application in larger datasets for further validation.

In [8], the authors present a study utilizing Wavelet Transform (WT) for processing Electroencephalogram (EEG) signals. They apply linear discriminant analysis (LDA) for feature selection and dimensionality reduction, utilizing the resulting two-dimensional features as a benchmark for classification via a Multi-Layer Perceptron (MLP) neural network. The proposed model achieves a high sensitivity, specificity, and accuracy of 100% for five classification problems. Comparative analysis with previous literature methods demonstrates the superiority of their approach in EEG signal classification and automated diagnosis. In [9], the authors introduce a novel EEG classification network, Separable EEGNet (S-EEGNet), leveraging Hilbert–Huang Transform (HHT) and a separable Convolutional Neural Network (CNN) with bilinear interpolation. HHT transforms EEG signals into a time-frequency representation, enhancing their description in the frequency domain.
The network combines depthwise and pointwise elements to extract feature maps, incorporating a bilinear interpolation method for flexible grid deformation based on local, dense, and adaptive EEG data characteristics. S-EEGNet conducts end-to-end learning of time and space dimensions, improving EEG classification accuracy. Experimental results on two EEG datasets demonstrate accuracy improvements of 3.6%, 1.15%, and 1.33% for motor imagery classification, and 89.91%, and 88.31% for emotion classification. In [10], the authors present a machine learning-based classifier designed to differentiate between Schizophrenia patients and healthy controls using features extracted from electroencephalograph (EEG) signals based on event-related potential (ERP). Utilizing an online EEG dataset with 81 subjects (32 healthy controls and 49 Schizophrenia patients), various features are extracted, preprocessed, and subjected to the random forest machine learning algorithm due to the dataset’s limited size. The study concludes that the classification accuracy can be enhanced by incorporating specific features extracted from EEG signals. In [11], the authors introduce an EEG classification approach for imagined speech with high accuracy and efficiency. Utilizing correntropy spectral density (CSD) matrices from EEG signals across different channels, the distances between these matrices serve as measures for imagined speech recognition. The simplicity and accuracy of Riemannian distance, a statistical method, are leveraged in the evaluation of channel selection and frequency band detection during imagined speech. The study, conducted on the "Kara One" database featuring EEG signals during imagined speech, demonstrates an average classification accuracy of 90.25% for all subjects, showcasing the efficiency and accuracy of the proposed method compared to other imagined speech classification approaches. In [12], the authors present a novel approach for the automatic detection of seizure activity in epileptic EEG signals. They employ a Long Short-Term Memory (LSTM) network for classification, utilizing Discrete Wavelet Transform (DWT) for noise removal and feature extraction, resulting in 20 eigenvalue features. Feature selection is optimized through correlation and P-value analysis, aiming to minimize LSTM trainable parameters while maintaining high accuracy. Comparative analysis demonstrates the superiority of the proposed method over other frameworks, including logistic regression (LR), support vector machine (SVM), K-nearest neighbor (K-NN), and decision tree (DT). In [13], the authors introduce a novel feature selection technique based on Rényi min-entropy to enhance the efficiency of brain–computer interface (BCI). In the context of multiple-class problems, where classification accuracy relies on effective feature selection from wavelet packet transformation (WPT) features extracted from electroencephalogram (EEG) signals, conventional methods use Shannon entropy and mutual information for feature selection. The proposed Rényi min-entropy-based approach is demonstrated to outperform these traditional methods in the classification of multiple EEG signals. The BCI competition-IV dataset, containing 4-class motor imagery EEG signals, is utilized for experimentation, showcasing the superiority of the proposed approach in multiple-class BCI scenarios. In [14], the authors conduct a comprehensive comparison of various functional neuroimaging techniques, highlighting the exceptional capabilities of Electroencephalogram (EEG) signals, including high temporal resolution, affordability, portability, and non-invasiveness. They explore different frequency bands associated with brain signals and provide a literature survey covering diverse applications of EEG signals in computer-aided technologies. These applications range from diagnosing neurological disorders like epilepsy and major depressive disorder to monitoring motor imagery, identity authentication, emotion recognition, sleep stage classification, eye state detection, and drowsiness monitoring. The survey includes a comparative analysis of EEG datasets, data acquisition methods, preprocessing and feature extraction techniques, classification models, and statistical tests. The paper also identifies research gaps and outlines future directions, encouraging further exploration in the field. The survey concludes with a brief overview of studies investigating the fusion of brain signals from multiple modalities. In [15], the authors introduce a novel approach, Pin-UTSVM, for the classification of Electroencephalogram (EEG) signals. They address the computational complexity of Support Vector Machines (SVMs) and twin support vector machines (TWSVM) by incorporating prior information through universum data in universum SVM (USVM) and universum twin SVM (UTSVM) models. The proposed Pin-UTSVM, utilizing the pinball loss function, enhances stability and noise resilience, maintaining computational efficiency. Universum data, specifically interictal EEG signals, is employed in the approach. Experimental results across various noise levels demonstrate the robustness of Pin-UTSVM, outperforming existing models in multiple feature extraction techniques. Statistical tests affirm the significant superiority of Pin-UTSVM over baseline models in EEG signal classification.

The literature review underscores the multifaceted landscape of EEG signal classification, revealing the dynamic interplay between advancements in machine learning and the intricate realm of neuroscience. Our approach uses DWT and statistical features and selects the best features using evolutionary algorithms for EEG signal classification.

III. Methodology

In our proposed methodology as shown in Figure 1, we initiate the process by taking EEG features as input, recognizing the significance of these features in capturing the intricate electrical activities of the brain. To enhance efficiency and eliminate redundancy, we employ Principal Component Analysis (PCA) to reduce the dimensionality of the feature space. This reduction not only streamlines the subsequent processes but also retains the most critical information for classification. Subsequently, we employ Discrete Wavelet Transform (DWT) and statistical feature extraction to delve deeper into the essence of the EEG signals, capturing both temporal and statistical characteristics. Recognizing the importance of selecting the most pertinent features, we incorporate evolutionary algorithms, specifically Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), to iteratively identify the optimal subset of features. Finally, employing a machine learning algorithm, such as Support Vector Machines (SVM)
or Decision Trees, we classify the EEG signals based on the refined and selected features, resulting in a comprehensive and effective approach for EEG signal classification.

A. Data

An EEG (Electroencephalogram) signal dataset typically consists of a series of numerical values corresponding to different channels or electrodes capturing electrical brain activity. In the provided snippet, X1 through X178 likely represent data points or readings from ten different channels or features within the EEG dataset. Each variable (X1, X2, ..., X178) corresponds to a specific measurement or attribute of the EEG signal, providing information about the electrical activity at different points on the scalp or brain regions. These values could represent voltage measurements, frequency components, or other relevant features extracted from the EEG signals during a specific time period, allowing for the analysis and classification of brain activity patterns.

B. Feature Reduction

The utilization of Principal Component Analysis (PCA) [16] in processing EEG signals serves as a strategic approach to mitigate redundant features within the dataset. PCA is employed to transform the original high-dimensional EEG signal dataset into a reduced-dimensional space while retaining essential information. By identifying and eliminating redundant features, PCA helps streamline the dataset, enhancing computational efficiency and simplifying subsequent analysis. This technique aids in focusing on the most significant patterns and variations in the EEG signals, contributing to a more refined and informative representation of the neural activity recorded by the EEG.

Principal Component Analysis (PCA) [16] is a statistical technique widely used in data analysis and dimensionality reduction. It serves as a valuable tool for transforming high-dimensional datasets into a lower-dimensional space while retaining essential information. The importance of PCA lies in its ability to simplify complex datasets, eliminate redundancies, and identify the most significant patterns or features. Mathematically, PCA [16] involves finding the eigenvectors and eigenvalues of the covariance matrix of the original dataset. The covariance matrix captures the relationships between different variables. Let X be the data matrix with each row representing an observation and each column representing a variable. The covariance matrix

\[ C = \frac{1}{n-1}(X - \bar{X})^T(X - \bar{X}) \]

, where n is the number of observations and \( \bar{X} \) is the mean vector. The next step is to find the eigenvectors \((v_1, v_2, ..., v_p)\) and eigenvalues \(\lambda_1, \lambda_2, ..., \lambda_p\) of \(C\). These
eigenvectors represent the directions of maximum variance in the original data, and the corresponding eigenvalues quantify the amount of variance along those directions.

The principal components are then obtained by selecting the top k eigenvectors corresponding to the k largest eigenvalues. The transformed dataset $Y$ is given by $Y = X.V_k$, where $V_k$ is the matrix containing the selected eigenvectors. In summary, PCA [15] simplifies datasets by transforming them into a space defined by their principal components, preserving the most critical information while reducing dimensionality. This process facilitates better visualization, noise reduction, and improved efficiency in subsequent data analysis tasks.

C. Feature Extraction

In the process of feature extraction, our approach considers a combination of statistical features and features obtained through Discrete Wavelet Transform (DWT), all derived from the reduced dimensionality features.

1) Statistical Features:
(a) Hjorth Features: Hjorth parameters are statistical measures used to analyze the time-domain characteristics of a signal. They are commonly employed in the analysis of electroencephalogram (EEG) signals. The three Hjorth parameters are:

- Activity: Activity represents the total energy or the variance of the signal. It is computed as the variance of the signal. Higher activity values indicate a higher level of overall signal energy.
- Mobility: Mobility measures the mean frequency or the average frequency at which the energy is distributed in the signal. It is computed as the ratio of the standard deviation of the first derivative of the signal to the standard deviation of the signal. Higher mobility values suggest a higher mean frequency and greater frequency dispersion in the signal.
- Complexity: Complexity is an indicator of the signal’s irregularity or waveform complexity. It is computed as the ratio of the mobility of the first derivative to the mobility of the signal. Higher complexity values indicate a more irregular and complex signal waveform.

These Hjorth parameters offer insights into different aspects of the signal, with Activity representing overall energy, Mobility capturing frequency characteristics, and Complexity describing the irregularity or complexity of the signal waveform.

(b) Kurtosis: Kurtosis is a statistical measure that assesses the tailedness and presence of outliers in the distribution of the EEG signal. Higher kurtosis values indicate heavier tails and a higher likelihood of outliers in the signal distribution.

(c) 2nd Diff Mean: This feature represents the mean of the second derivative of the EEG signal. It provides insights into the average rate of change or curvature of the signal.

(d) 2nd Diff Max: Denotes the maximum value of the second derivative of the EEG signal. It highlights the maximum rate of change or curvature in the signal.

(e) Coefficient of Variation: Measures the relative variability or dispersion of the EEG signal. A higher coefficient of variation indicates greater relative variability in the signal.

(f) Skewness: Indicates the asymmetry of the distribution of the EEG signal. Positive skewness suggests a longer tail on the right, while negative skewness suggests a longer tail on the left.

(g) 1st Difference Mean: Represents the mean of the first difference of the EEG signal. It offers insights into the average rate of change in the signal.

(h) 1st Difference Max: Denotes the maximum value of the first difference of the EEG signal. Higher variance indicates greater variability in the signal values.

(i) Variance: Measures the spread or dispersion of the EEG signal. Higher variance indicates greater variability in the signal values.

(j) Mean of Vertex to Vertex Slope: Represents the average slope between successive maxima and minima in the EEG signal. It characterizes the average steepness of the signal between consecutive extrema.

In summary, these statistical features provide diverse insights into different aspects of the EEG signal’s distribution, variability, and rate of change, contributing to a comprehensive analysis of the signal’s characteristics.

2) Discrete Wavelet Transform: The Discrete Wavelet Transform (DWT) [17] is a mathematical technique used for analyzing signals and images. It decomposes a signal into different frequency components, revealing both low and high-frequency information. DWT is widely employed in various applications, including signal processing, image compression, and denoising. Mathematically, the one-dimensional DWT for a discrete signal $x[n]$ is expressed as follows:

$$W(a, b) = \sum nX[n].\psi_{a,b}[n]$$

Here, $W(a, b)$ represents the wavelet coefficients after transformation. $a$ and $b$ are scaling and translation parameters. $\psi_{a,b}[n]$ is the wavelet function.

The two parameters, $a$ and $b$, control the scaling and translation of the wavelet function. The DWT is applied by successively dividing the signal into approximation (low-frequency) and detail (high-frequency) components through a process known as “downsampling” or “decimation.” For a two-dimensional signal (e.g., an image), the DWT can be extended similarly. The process involves applying the one-dimensional DWT separately to rows and columns. The key steps of the DWT algorithm involve a series of convolutions and downsamplings using the wavelet function. The result is a set of coefficients that represent different levels of detail and approximation of the input signal or image. DWT is known for its ability to capture both localized and global features of signals and images efficiently, making it a powerful tool in various signal processing applications.

3) Wavelet transform features: These features are computed based on the Discrete Wavelet Transform (DWT) [17] and provide statistical and informational measures for both
the approximate and detailed coefficients obtained through the wavelet transform.

(a) Wavelet Approximate Mean: It calculates the mean (average) of the approximate coefficients obtained through wavelet transform. Represents the central tendency of the approximate coefficients, providing insight into the overall energy distribution at the approximate level.

(b) Wavelet Approximate Std Deviation: Computes the standard deviation of the approximate coefficients. Measures the dispersion or spread of the approximate coefficients, indicating how much they deviate from the mean. Higher values suggest greater variability.

(c) Wavelet Detailed Mean: Calculates the mean of the detailed coefficients obtained through wavelet transform. Reflects the central tendency of the detailed coefficients, giving information about the overall energy distribution at the detailed level.

(d) Wavelet Detailed Std Deviation: Computes the standard deviation of the detailed coefficients. Indicates the variability or spread of the detailed coefficients around their mean. Higher values imply greater fluctuations.

(e) Wavelet Approximate Energy: Measures the energy of the approximate coefficients. Provides an indication of the overall strength or magnitude of the approximate coefficients, revealing the contribution of each coefficient to the signal.

(f) Wavelet Detailed Energy: Calculates the energy of the detailed coefficients. Represents the strength or magnitude of the detailed coefficients, indicating their contribution to capturing fine details in the signal.

(g) Wavelet Approximate Entropy: Computes the entropy of the approximate coefficients. Quantifies the unpredictability or disorder in the approximate coefficients, offering insights into the complexity of the signal.

(h) Wavelet Detailed Entropy: Measures the entropy of the detailed coefficients. Quantifies the unpredictability or disorder in the detailed coefficients, providing information about the complexity of fine details in the signal.

These features offer a comprehensive analysis of the statistical properties, energy distribution, and informational content at both the approximate and detailed levels of the wavelet-transformed signal.

The process involves the extraction of features from EEG signals, employing both statistical and wavelet transform coefficient methods. Statistical features encompass measures like mean, standard deviation, skewness, kurtosis, and various others, offering insights into the distribution and characteristics of the EEG signal. These features capture essential statistical properties, aiding in understanding the central tendency, variability, and shape of the EEG signal’s distribution. Utilizing wavelet transform involves decomposing the EEG signal into different frequency components using wavelets. Features are then derived from the coefficients obtained during this transformation. Wavelet transform captures both time and frequency information, making it suitable for analyzing signals with non-stationary characteristics, such as EEG signals. The calculated statistical features and wavelet transform coefficient features are integrated to form a set of final features. This integration enables a comprehensive representation of the EEG signal, combining information about its statistical properties and frequency characteristics.

The ultimate goal of this feature extraction process is to create a more enhanced and representative feature set that encapsulates crucial aspects of the EEG signals. These extracted features are often utilized as inputs for machine learning algorithms, contributing to tasks like EEG signal classification for various neurological applications. The combination of statistical features and wavelet transform coefficients ensures a thorough analysis of EEG signals, capturing both their statistical properties and frequency characteristics. This process lays the groundwork for subsequent tasks such as classification and interpretation in the realm of EEG signal analysis.

D. Feature Selection

The process involves the integration of features derived from both Discrete Wavelet Transform (DWT) and statistical analyses. These combined features serve as the input to sophisticated feature selection algorithms, namely Particle Swarm Optimization (PSO) and Genetic Algorithm (GA). This integration represents a comprehensive approach, leveraging the strengths of both DWT, which captures frequency information in the signal, and statistical features, which provide insights into its distribution and characteristics. By subjecting this amalgamated feature set to the optimization power of PSO and GA, the selection algorithms work synergistically to identify the most relevant and discriminative features, optimizing the input for subsequent stages, such as machine learning-based EEG signal classification.

1) Genetic Algorithm: Genetic Algorithms (GAs) [18] are optimization algorithms inspired by the process of natural selection and genetics. They are used to find approximate solutions to optimization and search problems. The algorithm works by evolving a population of potential solutions over multiple generations to improve their fitness for solving a particular problem.

   (a) Basic Concepts:

   - Population: A set of potential solutions to the problem, represented as individuals or chromosomes.
   - Chromosome: A representation of a potential solution, often as a binary string or a set of parameters.
   - Fitness Function: A function that evaluates the performance or quality of a solution. It assigns a fitness score to each individual in the population.
   - Selection: The process of choosing individuals from the population to form a new generation based on their fitness scores. Higher fitness individuals have a higher chance of being selected.
   - Crossover (Recombination): The process of combining genetic material of two parents to produce offspring. It mimics the crossover of genetic material in biological reproduction.
   - Mutation: A random change applied to some individuals in the population to introduce diversity.
Particle Swarm Optimization (PSO) [19] is a population-based optimization algorithm inspired by the social behavior of birds and fish. It simulates the social interaction and cooperation among individuals in a swarm to find the optimal solution in a search space. PSO is a b) Algorithm:

- Initialization: Generate an initial population of individuals.
- Evaluation: Evaluate the fitness of each individual using the fitness function.
- Selection: Select individuals based on their fitness scores. Higher fitness individuals have a higher probability of being chosen.
- Crossover (Recombination): Combine genetic material of selected parents to create offspring.
- Mutation: Introduce random changes in some individuals to maintain diversity.
- Replacement: Create a new population from parents and offspring.
- Termination: Repeat the process for multiple generations or until a stopping criterion is met.

(b) Algorithm Steps:

1. Initialization: Initialize particles with random positions and velocities.
2. Evaluation: Evaluate the fitness of each particle using the objective function.
3. Update Personal Best (PBest): Update the personal best position for each particle based on its current fitness.
4. Update Global Best (GBest): Update the global best position based on the fitness of particles in the entire swarm.
5. Update Velocity and Position: Update the velocity and position of each particle based on its previous velocity, personal best, and global best positions.
6. Termination: Repeat steps 2-5 for multiple iterations or until a stopping criterion is met.

(c) Mathematical Background:

- Representation (Chromosome): 
  \[ X = [x_1, x_2, ..., x_n] \]
  where \( x_i \) represents a gene or parameter.
- Fitness Function: \( F(X) \), where \( F \) evaluates the quality of the solution represented by \( X \).
- Selection: Probability of selection for an individual 
  \[ i : P(i) = \frac{F(i)}{\sum F(j)} \]
- Crossover (Recombination): Combine genetic material from two parents, e.g., single-point crossover:
  \[ X_{offspring} = [x_1, ..., x_k, ..., x_n] \]
  where \( k \) is a crossover point.
- Mutation: Randomly change genes, e.g.,
  \[ x_i = x_i + \text{mutation rate} \times \text{random number} \]

Genetic algorithms are versatile and applicable to various optimization problems, and their effectiveness often lies in tuning parameters such as mutation rate, crossover rate, and population size.

2) Particle Swarm Optimization: Particle Swarm Optimization (PSO) [19] is a population-based optimization algorithm inspired by the social behavior of birds and fish. It is used for solving optimization problems where the objective is to find the optimal solution in a search space. PSO simulates the social interaction and cooperation among individuals in a swarm to search for the best solution.

(a) Basic Concepts:

- Particle: An individual solution in the search space, represented by a position vector.
- Swarm: A set of particles that move through the search space to find the optimal solution.
- Position: A vector representing a solution in the search space.
- Velocity: The rate of change of a particle’s position.
- Acceleration Coefficients: \( c_1 \) and \( c_2 \) emphasize local exploitation. Acceleration Coefficients \( (c_1 \text{ and } c_2) \): They determine the impact of personal and global best positions on the particle’s movement.

(b) Algorithm Steps:

1. Initialization: Initialize particles with random positions and velocities.
2. Evaluation: Evaluate the fitness of each particle using the objective function.
3. Update Personal Best (PBest): Update the personal best position for each particle based on its current fitness.
4. Update Global Best (GBest): Update the global best position based on the fitness of particles in the entire swarm.
5. Update Velocity and Position: Update the velocity and position of each particle based on its previous velocity, personal best, and global best positions.
6. Termination: Repeat steps 2-5 for multiple iterations or until a stopping criterion is met.

(c) Mathematical Background:

- Representation (Particle): A particle \( i \) is represented by a position vector
  \[ X_i = [x_{i1}, x_{i2}, ..., x_{in}] \]
  in an \( n \)-dimensional search space.
- Velocity Update: The velocity of particle \( i \) is updated using the formula:
  \[ V_{i(t+1)} = w \cdot V_{i(t)} + c_1 \cdot r_1 \cdot (P_{best_i} - X_i) + c_2 \cdot r_2 \cdot (G_{best} - X_i) \]
  where \( w \) is the inertia weight, \( c_1 \) and \( c_2 \) are acceleration coefficients, and \( r_1 \) and \( r_2 \) are random numbers.
- Position Update: The position of particle \( i \) is updated using the formula:
  \[ X_{i(t+1)} = X_{i(t)} + V_{i(t+1)} \]

Inertia Weight \( (w) \): It controls the impact of the previous velocity on the new velocity. A higher \( w \) emphasizes global exploration, while a lower \( w \) emphasizes local exploitation. Acceleration Coefficients \((c_1 \text{ and } c_2)\): They determine the impact of personal and global best positions on the particle’s movement.

E. Classification

The chosen features, derived through the combined process of feature extraction and selection, are then fed into classification algorithms. In this stage, the primary objective is to employ machine learning techniques to categorize the EEG signals based on the discriminating features identified earlier. The classification algorithms, such as Support Vector Machines (SVM), Naive Bayes, Decision Trees, and Random Forest, utilize the selected features to distinguish and assign the EEG signals to different predefined classes. This final step completes the proposed methodology, where the integration of advanced feature extraction, selection, and classification techniques aims to enhance the accuracy and efficiency of EEG signal classification for a more nuanced understanding of neurological conditions.
1) Support Vector Machine (SVM): SVM is a supervised machine learning algorithm used for classification and regression tasks. It finds a hyperplane in an N-dimensional space that separates classes with a maximum margin. It is effective in high-dimensional spaces and capable of handling both linear and non-linear data. It is Robust against overfitting, effective in high-dimensional spaces.

2) Naive Bayes: Naive Bayes is a probabilistic classification algorithm based on Bayes’ theorem. It assumes independence among features, hence the term "naive." It is simple and computationally efficient. Performs well on text classification and spam filtering. The algorithm is fast and efficient, works well with categorical data.

3) k-Nearest Neighbors (KNN): KNN is a simple, instance-based learning algorithm. It classifies a new data point based on the majority class of its k nearest neighbors. No model training phase; it memorizes the entire dataset. Effective for small to medium-sized datasets. The algorithm is simple and intuitive, works well with locally clustered data. Cons: Computationally expensive for large datasets, sensitive to irrelevant features.

4) Random Forest: Random Forest is an ensemble learning method that constructs a multitude of decision trees during training. The final prediction is the mode of the predictions of individual trees (classification) or the average (regression). It combines multiple weak learners to create a strong learner. Reduces overfitting and improves accuracy. It handles high-dimensional data, robust to outliers.

We conducted a comprehensive comparison of all the algorithms by evaluating their performance using selected features obtained through evolutionary algorithms, namely Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). Additionally, we compared their performance with the utilization of all available features. This thorough evaluation allows us to assess the effectiveness of the feature selection process performed by GA and PSO in enhancing the classification accuracy of the algorithms compared to the scenario where all features are considered. The goal is to identify the most optimal combination of algorithms and features that yield the highest precision and recall in the classification of EEG signals.

IV. RESULTS

The initial dataset, sourced from the reference, comprises 5 distinct folders, each containing 100 files representing individual subjects. Each file records brain activity over a duration of 23.6 seconds, resulting in a time-series of 4097 data points. These data points represent EEG values. This results in 11500 pieces of information, with 178 data points per second. The data is further categorized into non-seizure and seizure activities. It undergoes division into training and testing datasets using k-fold cross-validation. The standardized database is then fed into a feature extraction algorithm based on the aforementioned approach. The accuracy of the classification model is calculated and compared.

PSO and Classification algorithms confusion matrix are shown in Figure 3 to Figure 6.
GA and Classification algorithms test set accuracy are shown in Figure 7 to Figure 10.

V. CONCLUSION AND FUTURE WORK

In conclusion, our investigation into the classification of EEG signals through the synergistic utilization of Discrete Wavelet Transform (DWT) for feature extraction and evolutionary algorithms for feature selection has yielded promising results. The comprehensive evaluation of various machine learning algorithms, including Support Vector Machines (SVM), Naive Bayes, Decision Trees, and Random Forest, has provided insights into their efficacy for accurate EEG signal classification. By addressing the intricate nature of neurological conditions, particularly epilepsy and seizures, our approach aims to enhance diagnostic capabilities. The successful implementation of genetic algorithms (GA) and Particle Swarm Optimization (PSO) in feature selection underscores their potential for optimizing the discriminative power of selected features, contributing to improved classification accuracy.

Building upon the foundations laid in this study, future work in the realm of EEG signal classification holds substantial potential for advancements. First and foremost, the exploration
of additional machine learning algorithms and their comparative analysis could provide further refinement in classification accuracy. Further research may delve into the integration of deep learning techniques, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs), to unravel more intricate patterns within EEG signals. Additionally, the consideration of diverse datasets encompassing a broader spectrum of neurological disorders would contribute to the generalizability of the classification model. Incorporating real-time applications and deploying the model in clinical settings could validate its effectiveness and pave the way for practical implementations in diagnostic procedures. Collaborations with medical practitioners and neuroscientists could provide valuable insights and ensure the translation of our findings into impactful clinical solutions. The continual evolution of machine learning and signal processing methodologies offers an exciting trajectory for future endeavors in EEG signal classification. By addressing the outlined avenues for future work, we aspire to contribute to the ongoing efforts to enhance our understanding of neurological conditions and improve diagnostic tools for the benefit of individuals affected by such disorders.

REFERENCES


Mayur Akewar is an AI researcher and Alumni of Department of Computer Science & Engineering, Shri Ramdeobaba College of Engineering & Management. He has total 12 years of academic and industrial experience. His background in AI spans around machine learning, deep learning and algorithms to deal with the problems related to computer vision including Hyperspectral Imaging (HSI) and natural language processing, and he has a proven track record of developing AI-based solutions, conducting research, or collaborating with organizations. He developed and designed many algorithms and trained and fine tuned many ML and DL models for applications of object detection, image, video, audio and text classification and generation.