

A Comprehensive Review of Feature Extraction and Classification Techniques in EEG Signal Analysis

Soha Savadi

Department of Computer Science, Faculty of Advanced Technologies, University of Mohaghegh Ardabili,
Namin, Iran

ABSTRACT

Brain-computer interface Electroencephalography (EEG) is a non-invasive technique widely used in neuroscience, clinical diagnosis, and brain-computer interface (BCI) applications. The accurate interpretation of EEG signals relies on robust signal processing, particularly in feature extraction and classification, which transform raw signals into meaningful representations for analysis and informed decision-making. This review presents a comprehensive overview of feature extraction techniques in the time, frequency, time-frequency, spatial, and nonlinear domains, as well as recent deep learning-based feature learning approaches. Furthermore, it examines classical and modern classification algorithms tailored to EEG analysis, covering both traditional machine learning and deep neural networks. The paper discusses current challenges, including low signal-to-noise ratio, inter-subject variability, and data scarcity. It explores promising future directions, including explainable AI, transfer learning, and real-time processing. By summarizing key advances and highlighting gaps in the literature, this review serves as a valuable resource for researchers and engineers working on EEG-based systems.

KEYWORDS

Electroencephalography, Brain-computer interface, Feature Extraction, Classification.

Date of Submission: 15-06-2025

Date of acceptance: 30-06-2025

I. INTRODUCTION

Electroencephalography (EEG) records electrical activity generated by neuronal firing in the cerebral cortex using electrodes placed on the scalp or directly on the cortex. Due to its high temporal resolution, cost-effectiveness, and non-invasiveness, EEG is extensively used in clinical and research contexts, including seizure detection, sleep analysis, cognitive workload monitoring, emotion recognition, and brain-computer interfaces (BCIs) [1,2].

However, EEG signals are inherently noisy, non-stationary, and susceptible to artifacts such as muscle movements, eye blinks, and power line interference. Therefore, advanced signal processing is essential for extracting relevant patterns that can be classified into meaningful cognitive or pathological states. Feature extraction is a crucial step in this pipeline, as it transforms raw signals into compact and discriminative representations. Practical features may capture dynamics in the time, frequency, spatial, or nonlinear domain, depending on the application [3,4].

Classification, the next crucial step, maps the extracted features to the desired labels, such as seizure versus non-seizure or left-hand versus right-hand motor imagery. Traditional classifiers such as Support Vector Machines (SVMs), Linear Discriminant Analysis (LDA), and k-Nearest Neighbors (k-NN) have been used extensively, while deep learning methods like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks are gaining popularity due to their automatic feature learning capability [5,6].

This review paper aims to provide a systematic overview of EEG feature extraction and classification methods, bridging classical approaches and contemporary deep learning techniques. We also identify open challenges in the field and highlight future research directions that can enhance the reliability and scalability of EEG-based systems.

II. EEG SIGNAL CHARACTERISTICS

Electroencephalography (EEG) captures the brain's electrical activity with high temporal resolution by measuring voltage fluctuations across the scalp or cortex. These signals, typically in the range of 0.5 to 100 Hz, reflect the summated post-synaptic potentials of pyramidal neurons in the cerebral cortex [7]. Despite its long-

standing clinical utility, analyzing EEG signals is complex due to several intrinsic characteristics that influence the choice of preprocessing, feature extraction, and classification strategies. Figure 1 shows a pipeline for EEG signal processing.

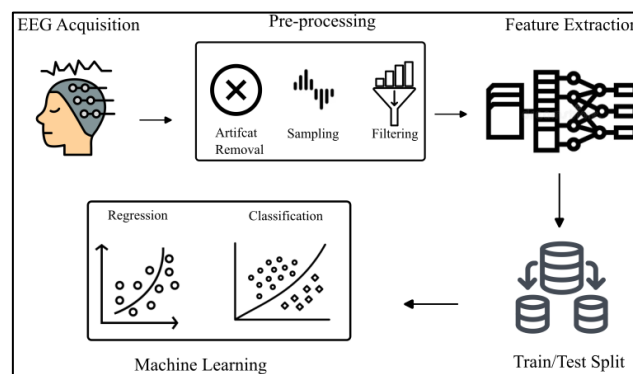


Figure 1: EEG signal processing pipeline.

2.1 Frequency Bands

EEG signals are commonly decomposed into standard frequency bands, each associated with distinct physiological and cognitive states:

- **Delta (0.5–4 Hz):** Observed during deep sleep and unconscious states.
- **Theta (4–8 Hz):** Linked to drowsiness, meditation, and early sleep stages.
- **Alpha (8–13 Hz):** Prominent during relaxed wakefulness, especially with closed eyes.
- **Beta (13–30 Hz):** Associated with active thinking, motor activity, and alertness.
- **Gamma (30–100 Hz):** Related to higher-order cognitive functions, such as attention and memory [8].

These frequency components form the basis for many feature extraction techniques, including bandpower estimation and spectral entropy analysis. Figure 2 shows signal traces for different frequency bands.

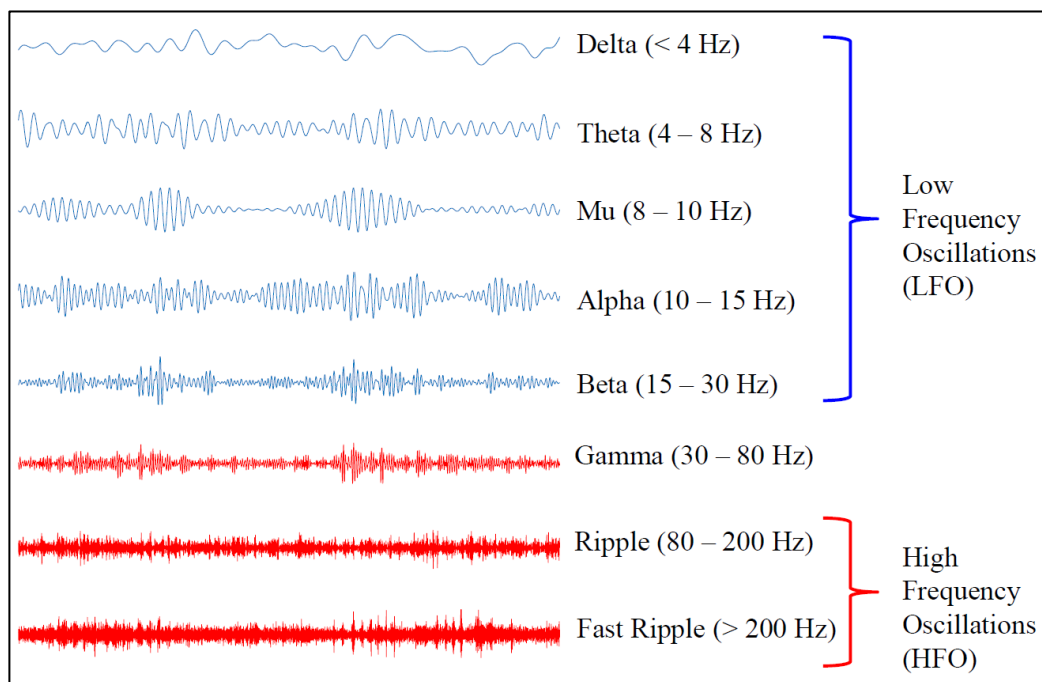


Figure 2: EEG frequency bands in the low and high frequency range.

2.2 Temporal and Spatial Resolution

One of EEG's main advantages is its millisecond-level temporal resolution, which makes it ideal for capturing transient cognitive or pathological events, such as epileptic spikes or event-related potentials (ERPs). However, due to volume conduction and the nature of scalp recordings, EEG has poor spatial resolution, often limiting its precision in localizing neural sources [9]. Advanced source localization algorithms and spatial filters

(e.g., Independent Component Analysis, Common Spatial Patterns) are employed to enhance spatial interpretability.

2.3 Non-Stationarity

EEG signals are non-stationary, meaning their statistical properties change over time. This variability may be induced by physiological factors (e.g., cognitive load, drowsiness) or external artifacts. As a result, classical signal processing methods (e.g., FFT) must often be complemented by time-frequency or adaptive strategies, such as Wavelet Transform or Empirical Mode Decomposition, to handle temporal fluctuations effectively [10].

2.4 Signal Amplitude and Noise

EEG amplitudes are typically in the range of 10–100 μ V and are easily contaminated by various sources of noise:

- **Physiological artifacts:** Eye movements (electrooculogram, EOG), muscle activity (electromyogram, EMG), heartbeats (electrocardiogram, ECG).
- **Environmental artifacts:** Power line interference (50/60 Hz), electrode movement, cable noise.

Artifact removal is often a prerequisite for reliable feature extraction and classification. Techniques such as regression, blind source separation (e.g., ICA), and filtering are widely used [11].

2.5 Inter Subject & Intra Subject Variability

EEG characteristics vary significantly between individuals (inter-subject) and within the same individual over time (intra-subject), due to factors such as head shape, electrode impedance, mental state, and fatigue [12]. These variabilities present a significant challenge for building generalizable models and often necessitate subject-specific calibration or domain adaptation techniques.

III. PREPROCESSING TECHNIQUES

EEG preprocessing is a fundamental step in signal analysis, aiming to remove unwanted noise and artifacts while preserving informative neural activity. Given the low amplitude and high susceptibility of EEG signals to both internal and external interferences, effective preprocessing is crucial to ensure the quality and reliability of downstream feature extraction and classification tasks. This section reviews the most common EEG preprocessing methods, including filtering, artifact removal, baseline correction, and normalization.

3.1 Filtering

Filtering is often the first step in EEG preprocessing. It eliminates irrelevant frequency components, such as DC offsets or power line noise, and focuses the analysis on frequency bands of interest.

- **Bandpass filters** (typically 0.5–70 Hz) are used to retain frequencies relevant to brain activity.
- **Notch filters** are applied at 50 or 60 Hz to suppress power line interference [13].
- **High-pass filters** (e.g., ≥ 0.5 Hz) remove slow drifts, while **low-pass filters** (e.g., ≤ 70 Hz) attenuate high-frequency noise [14].

Filtering must be designed carefully to avoid signal distortion or phase shifts, particularly when analyzing event-related potentials (ERPs) or transient features.

3.2 Artifact Removal

EEG signals are frequently contaminated by **non-neural artifacts**, which can significantly affect analysis and classification performance. Artifact removal techniques can be categorized into two broad groups: **manual methods** and **automated algorithms**.

Blind Source Separation (BSS): Independent Component Analysis (ICA) is one of the most widely used BSS techniques. It decomposes EEG into statistically independent components, allowing artifacts such as eye blinks or muscle activity to be identified and removed [15]. Other methods include Canonical Correlation Analysis (CCA) and Principal Component Analysis (PCA).

Regression-Based Methods: Regression techniques can subtract known artifact sources (e.g., EOG or EMG channels) from the EEG data to isolate the underlying neural activity. This method assumes that the artifact contribution is linearly correlated with reference channels [16].

Wavelet-Based Methods: Wavelet Transform enables multiresolution decomposition, making it suitable for detecting and eliminating artifacts that vary in time and frequency. Wavelet thresholding has been employed to remove high-amplitude transients, such as muscle bursts, without compromising the underlying neural signals [17].

3.3 Baseline Correction

Baseline correction is commonly applied in event-related analyses, where the mean voltage of a pre-stimulus window is subtracted from the post-stimulus period. This ensures that observed deflections are relative to a consistent reference level [18]. It helps eliminate slow drifts and improves the interpretability of ERP components.

3.4 Normalization and Standardization

Normalization reduces inter-channel and inter-subject variability by scaling features to a standard range, such as [0, 1] or zero-mean unit-variance. This is especially important when using machine learning algorithms that are sensitive to feature magnitude, such as SVMs or neural networks [19].

- **Z-score normalization:** Subtracts the mean and divides by the standard deviation.
- **Min-max normalization:** Scales features to a fixed interval, typically 0, 1.

Normalization ensures fair weighting of features during classification and improves convergence in optimization-based methods.

IV. FEATURE EXTRACTION METHODS

Feature extraction is a critical step in EEG analysis, aiming to transform raw data into a more manageable and informative form that enhances classification performance. Due to the multidimensional and non-stationary nature of EEG signals, a wide range of features have been proposed in various domains, including time, frequency, time-frequency, spatial, and nonlinear representations. More recently, deep learning techniques have also enabled the automatic learning of features directly from raw or minimally processed signals. This section categorizes and discusses the most widely used methods.

4.1 Time-Domain Features

Time-domain features are directly computed from the amplitude of EEG signals over time and are often computationally efficient.

- **Mean and Variance:** Represent central tendency and dispersion of the signal within a time window.
- **Zero-Crossing Rate:** Indicates the number of times the signal crosses the zero-voltage axis; can reflect signal roughness.
- **Hjorth Parameters:** Include *Activity* (signal power), *Mobility* (mean frequency), and *Complexity* (rate of frequency change) [20].
- **Skewness and Kurtosis:** Measure the asymmetry and peakedness of the signal distribution, respectively.
- **Peak-to-Peak Amplitude:** The difference between maximum and minimum values within a time window.

Time-domain features are handy for ERP-based paradigms and seizure detection.

4.2 Frequency-Domain Features

Frequency-domain features are extracted using spectral analysis techniques that decompose EEG into constituent frequency components.

- **Power Spectral Density (PSD):** Estimated using methods such as the Fast Fourier Transform (FFT) or Welch's method, it reflects the power distribution across frequency bands [21].
- **Band Power:** Average power in specific bands (delta, theta, alpha, beta, gamma), often used in sleep staging, BCI, and attention monitoring.
- **Spectral Entropy:** Measures the spectral complexity; higher entropy implies more uniform power distribution.
- **Spectral Edge Frequency (SEF):** Frequency below which a certain percentage (e.g., 95%) of total spectral power is contained.

Frequency-based features are robust and widely used in both clinical and cognitive EEG applications.

4.3 Time-Frequency Features

Time-frequency representations are particularly suitable for capturing the non-stationary nature of EEG signals.

- **Short-Time Fourier Transform (STFT):** Applies FFT over short overlapping windows to capture localized frequency changes.
- **Wavelet Transform (WT):** Decomposes the signal into multi-resolution components using wavelet basis functions, enabling the analysis of transient patterns, such as epileptic spikes or sleep spindles [22].
- **Hilbert-Huang Transform (HHT):** Combines Empirical Mode Decomposition (EMD) and Hilbert Transform to extract instantaneous frequencies from nonlinear and non-stationary data [23].

Wavelet-based features have been highly successful in seizure detection, motor imagery classification, and emotion recognition. Figure 3 shows a time-frequency map using the wavelet transform.

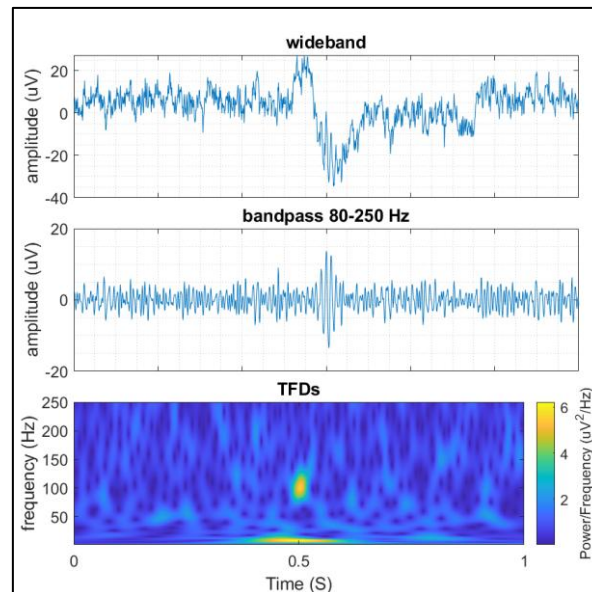


Figure 3: An example time-frequency map.

4.4 Spatial Features

Spatial filtering techniques improve the signal-to-noise ratio by enhancing class-discriminative activity across multiple EEG channels.

- **Common Spatial Patterns (CSP):** Maximizes variance for one class while minimizing it for another; widely used in motor imagery BCI [24].
- **Independent Component Analysis (ICA):** Separates EEG into statistically independent sources; also used for artifact removal.
- **Surface Laplacian:** Computes the second spatial derivative to enhance local brain activity and suppress volume-conducted signals.

These techniques exploit the spatial structure of EEG data, especially in multichannel recordings.

4.5 Nonlinear and Entropy Based Features

EEG signals exhibit chaotic dynamics and nonlinear behavior, which can be quantified using complexity measures.

- **Approximate Entropy (ApEn) & Sample Entropy (SampEn):** Quantify regularity and unpredictability in a time series.
- **Permutation Entropy:** Captures signal complexity based on ordinal patterns.
- **Fractal Dimension & Hurst Exponent:** Describe the self-similarity and long-range dependence in the signal [25].
- **Lyapunov Exponent:** Measures sensitivity to initial conditions, often used in seizure prediction.

Nonlinear features are compelling in pathological EEG classification tasks.

4.6 Deep Learning-Based Feature Learning

Deep learning models can automatically learn hierarchical features from raw EEG data, removing the need for handcrafted features.

- **Convolutional Neural Networks (CNNs):** Learn spatial and temporal filters directly from multichannel EEG inputs, making them effective in motor imagery, ERP, and seizure detection [26,27].
- **Recurrent Neural Networks (RNNs) and LSTM:** Capture temporal dependencies and long-term patterns in EEG.
- **Transformers:** Recent architectures that use self-attention to model long-range dependencies without recurrence.
- **Autoencoders:** Used for unsupervised representation learning and dimensionality reduction.

These methods require large datasets and significant computational resources but can outperform traditional pipelines when properly trained.

V. CLASSIFICATION

Once relevant features have been extracted from EEG signals, the next step is to assign them to predefined classes using appropriate classification algorithms. Classification is essential in various EEG-based applications, including seizure detection, brain-computer interface control, cognitive state monitoring, and sleep

stage classification. Depending on the task, both traditional machine learning methods and modern deep learning approaches are used. This section categorizes the main EEG classification strategies into conventional machine learning, ensemble methods, deep learning classifiers, and evaluation metrics.

5.1 Traditional Machine Learning Algorithms

Traditional classifiers have been extensively used for EEG classification due to their interpretability and effectiveness on relatively small datasets.

- **Support Vector Machine (SVM):** SVM is widely used due to its robustness in high-dimensional feature spaces and its ability to construct optimal hyperplanes. Kernel functions (linear, polynomial, radial basis function) can be employed to model nonlinear relationships in EEG data [28].
- **k-Nearest Neighbors (k-NN):** k-NN classifies a sample based on the majority class among its k-nearest neighbors in the feature space. It is non-parametric and straightforward, but can be sensitive to noise and computationally expensive with large datasets [29].
- **Linear and Quadratic Discriminant Analysis (LDA/QDA):** LDA assumes Gaussian-distributed classes with equal covariances and is popular in motor imagery BCI due to its computational efficiency. QDA is a nonlinear extension that allows each class to have its covariance matrix [30].
- **Naïve Bayes Classifier:** This probabilistic classifier assumes feature independence given the class label. While this assumption is often violated in EEG data, Naïve Bayes can still provide competitive performance in some contexts [31].
- **Decision Trees:** These create interpretable rules based on feature thresholds, but they are prone to overfitting. Regularization or pruning is typically necessary to avoid this issue.

5.2 Ensemble Learning Methods

Ensemble classifiers combine the predictions of multiple base learners to improve robustness and accuracy.

- **Random Forests (RF):** Random Forest is an ensemble of decision trees trained on random subsets of data and features. It reduces overfitting and increases classification stability [32].
- **AdaBoost and Gradient Boosting:** These methods iteratively combine weak classifiers to form a strong classifier. AdaBoost adapts to errors by reweighting instances, while Gradient Boosting uses gradient descent to minimize loss functions [33].

Ensemble techniques are particularly beneficial in EEG applications with complex decision boundaries or noisy features.

5.3 Deep Learning Classifiers

With the increasing availability of large EEG datasets and computational power, deep learning models have become prominent due to their ability to learn discriminative features directly from raw or minimally processed EEG data.

- **Convolutional Neural Networks (CNNs):** CNNs extract spatial and temporal patterns through convolutional filters. They have been widely used for motor imagery, sleep stage classification, and seizure detection [34]. Models such as EEGNet have demonstrated high accuracy with a relatively small number of parameters.
- **Recurrent Neural Networks (RNNs) and LSTMs:** RNNs are capable of modeling temporal dependencies in EEG signals. Long Short-Term Memory (LSTM) units address the vanishing gradient problem, making them suitable for capturing long-range dependencies [35].
- **Transformer-Based Models:** Recently, attention-based architectures such as transformers have been adapted for EEG analysis, enabling the model to focus on the most informative parts of the signal across time and space [36].
- **Hybrid Models:** Combining CNNs with LSTMs or attention mechanisms can further enhance classification by simultaneously leveraging spatial and temporal dependencies.

5.4 Evaluation Metrics

EEG classification performance is commonly evaluated using several metrics:

- **Accuracy (ACC):** Proportion of correctly classified samples.
- **Precision, Recall, F1-Score:** Useful for imbalanced datasets.
- **Area Under Curve (AUC):** Indicates classifier performance across decision thresholds.
- **Confusion Matrix:** Summarizes true vs. predicted class distributions.
- **Cross-Validation:** Techniques such as k-fold or leave-one-subject-out are often used to ensure generalization.

Choosing appropriate metrics is essential, especially in clinical applications where misclassifications can have serious consequences.

VI. CONCLUSION

Electroencephalography (EEG) remains one of the most widely used tools in neuroscience and clinical diagnostics due to its non-invasive nature and excellent temporal resolution. However, transforming raw EEG signals into meaningful information suitable for decision-making involves a complex pipeline of preprocessing, feature extraction, and classification.

In this review, we provided a structured and comprehensive overview of EEG feature extraction methods across time, frequency, time-frequency, spatial, and nonlinear domains. We also examined emerging deep learning-based approaches that automate feature learning from raw or minimally preprocessed EEG signals. Furthermore, we analyzed a broad spectrum of classification techniques, including traditional machine learning algorithms such as SVM, LDA, and Random Forest, as well as advanced deep neural networks, including CNNs, LSTMs, and transformers.

Each method brings specific advantages and limitations depending on the application, data availability, and computational constraints. For instance, handcrafted features remain useful in domains with limited data, while deep learning approaches excel when large, annotated datasets are available. Despite numerous advancements, several challenges persist, including inter-subject variability, low signal-to-noise ratio, and the need for robust real-time performance.

Future research should focus on developing generalizable and interpretable models, leveraging techniques such as transfer learning, domain adaptation, and explainable artificial intelligence. The integration of multimodal biosignals, edge computing, and privacy-preserving learning frameworks, such as federated learning, also holds significant promise.

Overall, the synergy between advanced signal processing techniques and powerful classification algorithms will continue to play a pivotal role in expanding the scope and impact of EEG-based technologies across clinical, cognitive, and human-computer interaction domains.

REFERENCES

- [1] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in Biology and Medicine*, vol. 100, pp. 270–278, Sep. 2018.
- [2] M. Mohammadpour, C. Kapeller, M. Korostenskaja, L. Schreiner, J. Scharinger, and C. Guger, "Seizure onset zone classification of intracranial EEG signals from epilepsy patients," in *Proc. 46th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Orlando, FL, USA, Jul. 2024, pp. 1–5.
- [3] S. Sanei and J. A. Chambers, *EEG Signal Processing*, Wiley, 2007.
- [4] M. Mohammadpour, C. Kapeller, K. Kamada, J. Scharinger, M. Schwarzgruber, M. Korostenskaja, and C. Guger, "Online detection of epileptic spikes for use in epilepsy monitoring," in *Proc. 9th Graz Brain-Computer Interface Conf.*, Graz, Austria, 2024.
- [5] G. Farjamnia, M. Z. Gashti, H. Barangi, and Y. S. Gasimov, "The study of support vector machine to classify the medical data", *IJCSNS International Journal of Computer Science and Network Security*, vol. 17, no. 12, pp. 145–150, 2017.
- [6] M. Z. Gashti, "A modified model based on flower pollination algorithm and k-nearest neighbor for diagnosing diseases", *IIUM Engineering Journal*, vol. 19, no. 1, pp. 144–157, 2018.
- [7] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [8] A. Delorme and S. Makeig, "EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [9] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd ed., Oxford University Press, 2006.
- [10] R. Q. Quiroga, "Quantitative analysis of EEG signals: time-frequency methods and chaos theory," *Journal of Neuroscience Methods*, vol. 94, no. 1, pp. 121–126, 1999.
- [11] T. K. Das and P. Puthankattil, "Wavelet-based ICA algorithm for artifact removal in EEG," *Neurocomputing*, vol. 146, pp. 449–455, Dec. 2014.
- [12] R. Chavarriaga et al., "Toward a hybrid BCI: combining brain and physiological signals to support intention recognition," *EPJ Nonlinear Biomedical Physics*, vol. 1, no. 1, pp. 1–11, 2013.
- [13] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publishing, 1997.
- [14] J. G. Webster (Ed.), *Medical Instrumentation: Application and Design*, 4th ed., Wiley, 2009.
- [15] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *NeuroImage*, vol. 34, no. 4, pp. 1443–1449, 2007.
- [16] M. Fatourehchi, A. Bashashati, R. K. Ward, and G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clinical Neurophysiology*, vol. 118, no. 3, pp. 480–494, 2007.
- [17] L. Zhang, P. Xu, D. Guo, R. Zhang, and D. Yao, "Robust removal of muscle artifact from single-channel EEG based on EEMD and ICA," *PLoS ONE*, vol. 8, no. 6, e77332, 2013.
- [18] S. J. Luck, *An Introduction to the Event-Related Potential Technique*, 2nd ed., MIT Press, 2014.
- [19] G. Li, D. Jiang, Z. Zhang, X. Cai, and C. Jin, "EEG classification based on entropy and SVM for epileptic seizure detection," *Journal of Electrical and Computer Engineering*, vol. 2019, Article ID 8519568, 2019.
- [20] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.
- [21] J. R. Hughes and E. G. John, "Power spectrum analysis of EEG data: a review," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 6, pp. 565–574, 1970.
- [22] M. Mohammadpour, C. Kapeller, M. Korostenskaja, J. Scharinger, and C. Guger, "Clustering epileptic spike waveforms for predicting seizure onset zone," in *Proc. 31st Nat. and 9th Int. Iranian Conf. Biomed. Eng. (ICBME)*, Tehran, Iran, Nov. 2024, pp. 1–5.
- [23] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Royal Soc. A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

- [24] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, Dec. 2000.
- [25] J. P. Zbilut and C. L. Webber, "Embeddings and delays as derived from quantification of recurrence plots," *Physics Letters A*, vol. 171, no. 3–4, pp. 199–203, 1992.
- [26] V. Lawhern, D. Solon, N. Waytowich, S. Gordon, C. Hung, and B. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, 2018.
- [27] M. Z. Gashti, "Detection of Spam Email by Combining Harmony Search Algorithm and Decision Tree", *Engineering, Technology & Applied Science Research (ETASR)*, vol. 7, no. 3, pp. 1713–1718, Jun. 2017.
- [28] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.
- [29] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [30] R. Blankertz, G. Dornhege, M. Krauledat, K. R. Müller, and G. Curio, "The non-invasive Berlin Brain–Computer Interface: fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.
- [31] A. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 2, pp. R1–R13, 2007.
- [32] M. Z. Gashti, "A novel hybrid support vector machine with decision tree for data classification," *International Journal of Advanced and Applied Sciences (IJAAS)*, vol. 4, no. 9, pp. 138–143, 2017.
- [33] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [34] V. Lawhern, D. Solon, N. Waytowich, S. Gordon, C. Hung, and B. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, 2018.
- [35] H. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," in *Proc. Int. Conf. on Learning Representations (ICLR)*, 2016.
- [36] Y. Xie, J. Xu, and Y. Zhang, "An attention-based deep learning framework for EEG classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1327–1336, 2021.