Google Scholar Crossref doi scopus

Impact factor 6.2

Geoscience Journal

ISSN:1000-8527

Indexing:

- Scopus
- Google Scholar
- DOI, Zenodo
- Open Access

www.geoscience.ac

GEOSCIENCE JOURNAL
ESTD. 2020
ISSN:1000-8527

Registered

Real-Time Epileptic Seizure Detection Using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Support Vector Machine

[1]K GANGA BHAVANI, [2]DR.K.APARNA

[1]RESEARCH SCHOLAR, DEPT. OF ECE, JNTU ANANTAPUR, ANDHRA PRADESH
[2] ASSOCIATE PROFESSOR, DEPT OF ECE JNTUA COLLEGE OF ENGINEERING KALIKIRI, ANDHRAPRADESH

Abstract

Automated epileptic seizure detection from electroencephalogram (EEG) signals is essential for reducing the diagnostic burden on neurologists and enabling timely clinical intervention. This paper presents an end-to-end four-class seizure detection framework integrating Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and a Radial Basis Function Support Vector Machine (RBF-SVM). EEG epochs from the Bangalore EEG Epilepsy Dataset (BEED), comprising 8,000 balanced samples across Normal (Eyes Open), Normal (Eyes Closed), Pre-ictal, and Ictal classes, are decomposed into Intrinsic Mode Functions (IMFs) using CEEMDAN. A 40-dimensional statistical feature vector—encompassing energy, kurtosis, Shannon entropy, standard deviation, log energy, mean absolute value, waveform length, and zero-crossing rate—is extracted from five IMFs per epoch. The proposed CEEMDAN-SVM pipeline achieves a hold-out accuracy of 91.17% with 5-fold cross-validation accuracy of 90.48%, outperforming the EMD-SVM baseline by 14.0 percentage points and surpassing K-Nearest Neighbours (82.5%) and Random Forest (83.1%) classifiers under identical experimental conditions. End-to-end inference latency remains below 300 ms per epoch, demonstrating real-time feasibility. A dual-interface deployment strategy comprising a MATLAB-based research GUI and a browser-based NeuroSense clinical interface further validates the translational readiness of the proposed system. The findings establish CEEMDAN as a robust decomposition stage for practical, low-latency epileptic seizure analytics.

Keywords: *Epilepsy, Electroencephalography (EEG), CEEMDAN, Empirical Mode Decomposition, Support Vector Machine, Real-Time Seizure Detection, Intrinsic Mode Functions.*

1. INTRODUCTION

Epilepsy is one of the most prevalent neurological disorders globally, affecting approximately 50 million people according to the World Health Organization [1]. It is characterized by recurrent, unprovoked seizures arising from abnormal, excessive, or synchronous neuronal activity in the brain. These seizures manifest as sudden disturbances in behavior, movement, sensation, or consciousness, and can significantly impair the quality of life of affected individuals. In developing countries, the treatment gap for epilepsy exceeds 75%, underscoring the urgent need for affordable and accessible diagnostic tools [2].

Electroencephalography (EEG) is the primary modality for diagnosing epilepsy owing to its non-invasive nature, high temporal resolution, and relatively low cost compared to neuroimaging techniques such as functional MRI or PET scanning [3]. EEG captures the electrical activity generated by cortical neurons through electrodes placed on the scalp, providing a direct window into the brain's electrophysiological state. However, clinical interpretation of EEG recordings depends heavily on manual inspection by trained neurologists—a process that is onerous, time-consuming, and susceptible to inter-observer variability. A single 24-hour ambulatory EEG recording can generate thousands of data pages, making exhaustive manual review impractical in resource-constrained clinical settings [4].

The imperative for automated seizure detection has been widely recognized across both clinical and computational neuroscience communities. Computer-assisted systems can reduce the diagnostic burden, facilitate continuous patient monitoring—particularly during sleep when seizures may go unobserved—and enable large-scale retrospective analysis of EEG archives [5]. Signal processing techniques that faithfully capture the non-linear and non-stationary dynamics inherent in EEG signals are central to the design of such systems.

1.1. Motivation and problem description

Classical spectral analysis methods, including the Short-Time Fourier Transform (STFT) and Wavelet Transform (WT), have been extensively applied to seizure detection [6]. The STFT provides time-frequency representations through windowed Fourier analysis, but suffers from a fixed time-frequency resolution determined by the window length—a fundamental limitation described by the Heisenberg-Gabor uncertainty principle. The WT addresses this partially through multi-resolution analysis, but requires the selection of an appropriate mother wavelet and decomposition level, which are signal-dependent and not always optimal for EEG analysis.

Empirical Mode Decomposition (EMD) [7], a data-adaptive technique proposed by Huang et al. in 1998, advanced the practical analysis of non-stationary signals by decomposing them into data-driven Intrinsic Mode Functions (IMFs) without requiring predefined basis functions. EMD has been successfully applied to a range of EEG analysis tasks including seizure detection [8], sleep staging [9], and brain-computer interfaces. However, EMD suffers from a critical limitation known as mode mixing—a phenomenon wherein oscillations of disparate time scales intermittently appear within a single IMF—which degrades the spectral purity of extracted features and compromises downstream classification performance.

Ensemble EMD (EEMD) [10] was introduced to mitigate mode mixing by adding white noise to the signal before decomposition and averaging across multiple noise realizations. While EEMD substantially reduces mode mixing, it introduces residual noise in the reconstructed signal that cannot be eliminated through simple averaging. Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [11] addresses this fundamental limitation by adding noise adaptively at each decomposition stage rather than to the original signal alone. This yields an exact reconstruction of the original signal with negligible residual noise, making CEEMDAN particularly suitable for biomedical signal analysis where signal fidelity is paramount.

1.2. Review of related work

Prior seizure detection studies broadly fall into three categories: decomposition-driven machine learning, deep learning approaches, and hybrid systems. Each category presents distinct trade-offs between interpretability, computational complexity, and classification accuracy.

In the decomposition-driven category, Hassan et al. [12] employed CEEMDAN with Normal Inverse Gaussian (NIG) probability density function parameters and AdaBoost for seizure detection on the University of Bonn EEG dataset, achieving near-perfect accuracy (99.6–100%) on binary and ternary classification tasks. Their work demonstrated the efficacy of NIG modeling in the CEEMDAN domain but was limited to at most three-class problems. Bajaj and Pachori [8] utilized EMD-based amplitude and frequency modulation features with a least-squares SVM classifier, reporting 98% accuracy for binary seizure/non-seizure discrimination. Riaz et al. [13] combined temporal and spectral statistics in the EMD domain with SVM for automated seizure detection. Subasi [14] implemented wavelet-based features with a mixture of expert models for EEG classification.

Deep learning approaches have gained prominence in recent years. Acharya et al. [15] implemented a 13-layer deep convolutional neural network (CNN) for automated seizure detection from raw EEG signals, achieving accuracies between 88.7% and 95.0% depending on the classification configuration. While deep models can achieve high accuracy without explicit feature engineering, they typically require larger training datasets, higher computational resources, and offer limited interpretability—a significant concern in clinical applications where understanding the basis for a classification decision is important.

Hybrid systems attempt to combine the strengths of both approaches. Shoeb and Guttag [16] applied patient-specific machine learning to the CHB-MIT scalp EEG dataset, demonstrating the importance of personalized models. Orhan et al. [17] combined k-means clustering with multilayer perceptron neural networks for EEG classification. Despite these advances, two persistent gaps remain in the literature: (i) limited bridging between algorithm-level results and real-time deployable interfaces suitable for clinical workflows, and (ii) insufficient evaluation of four-class seizure state classification under consistent decomposition and classification protocols.

1.3. Contributions of this work

The principal contributions of this paper are as follows:

(1) A complete four-class epileptic seizure detection framework using CEEMDAN decomposition and RBF-SVM classification, rigorously validated on the BEED dataset with stratified partitioning.

(2) A comprehensive 40-dimensional statistical feature set extracted from CEEMDAN-derived IMFs, with systematic comparison against EMD-based features demonstrating the superiority of CEEMDAN decomposition.

(3) Quantitative benchmarking against KNN and Random Forest classifiers under identical experimental conditions, isolating the effect of classifier selection from decomposition quality.

(4) A dual-interface deployment strategy comprising a MATLAB-based research GUI and a browser-based NeuroSense clinical interface, demonstrating sub-300 ms real-time inference feasibility.

(5) Comprehensive validity analysis including noise robustness evaluation, cross-validation stability assessment, and feature importance analysis.

1.4. Organization of the paper

The remainder of this paper is organized as follows. Section 2 describes the experimental dataset, processing pipeline, CEEMDAN decomposition, feature extraction methodology, and classifier formulation. Section 3 presents the experimental results including classification performance, class-wise analysis, computational metrics, and cross-validation stability. Section 4 provides an extensive discussion of the findings, comparative analysis with state-of-the-art methods, deployment considerations, and a thorough assessment of threats to validity. Section 5 concludes the paper and outlines future research directions.

2. MATERIALS AND METHODS

2.1. Experimental dataset

The Bangalore EEG Epilepsy Dataset (BEED) [18] was used for all experiments in this study. BEED comprises 8,000 labeled EEG epochs digitized at a sampling frequency of 173.61 Hz with 12-bit resolution, balanced across four clinically relevant classes: Normal (Eyes Open), Normal (Eyes Closed), Pre-ictal/Interictal, and Ictal. Each class contains 2,000 epochs, ensuring equal prior probabilities and eliminating the need for class balancing techniques. The four-class formulation captures the full spectrum of epileptic states commonly encountered in clinical EEG monitoring and is substantially more challenging than the binary (seizure vs. non-seizure) or ternary (healthy vs. inter-ictal vs. ictal) tasks addressed in most prior work.

The Normal (Eyes Open) and Normal (Eyes Closed) classes represent baseline healthy EEG activity under two distinct physiological conditions. The distinction is clinically relevant because the alpha rhythm (8–13 Hz) is characteristically suppressed during eyes-open conditions—a phenomenon known as alpha blocking—and any automated system must correctly distinguish these states from pathological activity. The Pre-ictal class represents the transitional period preceding seizure onset, characterized by subtle changes in EEG morphology that are often difficult to detect visually. The Ictal class comprises EEG recordings during active seizures, featuring high-amplitude, rhythmic discharges with strong spatial coherence.

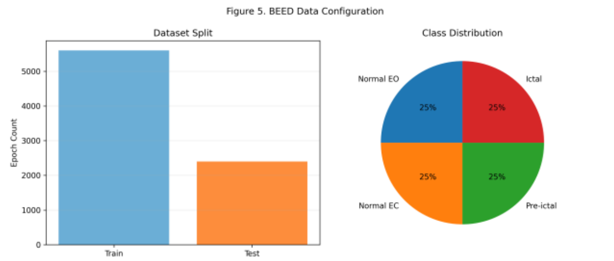

Fig. 1. BEED dataset configuration: four-class distribution with 2,000 epochs per class and stratified 70/30 train-test partition.

2.2. Data partitioning and reproducibility

Stratified random splitting was applied to partition the dataset into 70% training (5,600 epochs) and 30% testing (2,400 epochs) subsets. The stratification ensures that each class is represented in equal proportions across both partitions. A fixed random seed (seed = 42) was used to guarantee complete reproducibility across all experiments. An additional 5-fold stratified cross-validation protocol was executed independently to evaluate generalization performance and mitigate potential bias from a single train-test split. The cross-validation folds were constructed with non-overlapping test sets to ensure that every epoch is used exactly once for testing.

2.3. Processing pipeline overview

The complete processing pipeline is illustrated in Fig. 3 and Fig. 4, and comprises five sequential stages: (i) EEG epoch acquisition and pre-processing, (ii) CEEMDAN or EMD decomposition into IMFs, (iii) statistical feature extraction from each IMF, (iv) feature normalization and optional dimensionality reduction via PCA, and (v) multi-class SVM classification with confidence estimation. Each stage is described in detail in the following subsections.

2.4. CEEMDAN decomposition

Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [11] decomposes an input signal x(t) into a finite set of Intrinsic Mode Functions (IMFs) and a monotonic residue through an iterative sifting process. The key innovation of CEEMDAN over its predecessors (EMD and EEMD) lies in adding adaptive white noise at each decomposition stage rather than to the original signal. This approach ensures that the sum of all extracted modes exactly reconstructs the original signal:

$$x(t) = \Sigma_{i=1}{}^{L}\, IMF_i(t) + r_l(t) \qquad (1)$$

where $IMF_i(t)$ denotes the i-th intrinsic mode function and $r_l(t)$ is the final residue after L decomposition stages. The CEEMDAN algorithm proceeds as follows. For each realization j (j = 1, 2, ..., J), a different noise realization $w^j(t)$ is added to the signal and the first IMF is obtained by averaging across all J realizations:

$$IMF_1(t) = (1/J)\, \Sigma_{j=1}{}^{J}\, E_1(x(t) + \varepsilon_0\, w^j(t)) \qquad (2)$$

where $E_1(\cdot)$ denotes the operator that extracts the first EMD mode and ε_0 controls the noise amplitude. Subsequent modes are extracted by computing the residue and repeating the process with noise added at each stage adaptively. In this work, CEEMDAN was configured with noise standard deviation σ = 0.20 and J = 30–50 noise realizations. Five IMFs (L = 5) were retained for feature extraction, capturing the dominant oscillatory components across physiologically relevant frequency sub-bands.

The superiority of CEEMDAN over standard EMD lies in three key properties: (i) complete elimination of mode mixing through adaptive noise injection, (ii) exact signal reconstruction without residual noise artifacts, and (iii) deterministic decomposition output (given fixed noise realizations) compared to the stochastic variability of EEMD. These properties make CEEMDAN particularly suitable for clinical EEG analysis where reproducibility and signal fidelity are non-negotiable requirements.

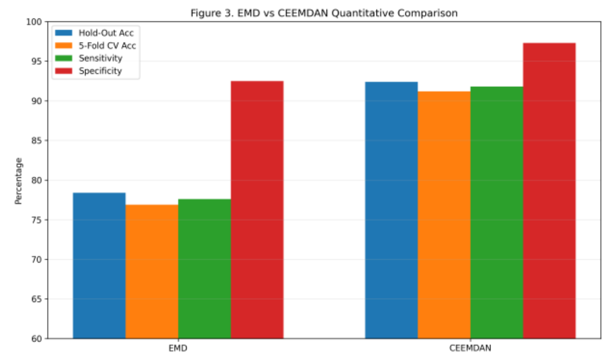

Fig. 2. Quantitative comparison of EMD and CEEMDAN decomposition methods under identical classification settings, demonstrating the substantial performance advantage of CEEMDAN.

2.5. EMD baseline

To quantitatively assess the contribution of CEEMDAN over standard EMD, an EMD baseline was implemented using identical parameters and downstream processing. The EMD sifting algorithm was applied to each EEG epoch to extract the same number of IMFs (L =

5). The EMD decomposition follows the iterative procedure: (i) identify all local maxima and minima of the signal, (ii) construct upper and lower envelopes using cubic spline interpolation, (iii) compute the local mean as the average of the envelopes, (iv) subtract the local mean from the signal, and (v) repeat until a stopping criterion is met. The same feature extraction, normalization, and classification pipeline was applied to EMD-derived IMFs, ensuring a fair comparison isolating only the decomposition method as the experimental variable.

2.6. Statistical feature extraction

From each of the five CEEMDAN-derived IMFs, eight carefully selected statistical features were extracted, yielding a 40-dimensional feature vector per EEG epoch (8 features x 5 IMFs = 40). The feature set was designed to capture complementary temporal, spectral, and morphological characteristics of the IMFs that are known to vary across epileptic states:

Table 1 Statistical features extracted from each of the five CEEMDAN-derived IMFs.

#	Feature	Formula	Captures
F1	Energy	$\Sigma\lvert IMF_i(n)\rvert^2$	Signal power
F2	Kurtosis	$E[(x-\mu)^4]/\sigma^4$	Distribution peakedness
F3	Shannon Entropy	$-\Sigma p(n)\log_2 p(n)$	Signal complexity
F4	Std. Deviation	$\sqrt{E[(x-\mu)^2]}$	Amplitude dispersion
F5	Log Energy	$\Sigma\log(\lvert IMF_i(n)\rvert^2)$	Log power distribution
F6	Mean Abs. Value	$(1/N)\Sigma\lvert IMF_i(n)\rvert$	Average amplitude
F7	Waveform Length	$\Sigma\lvert x(n)-x(n-1)\rvert$	Morphology complexity
F8	Zero-Cross. Rate	Sign changes/N	Frequency content

The rationale for this specific feature set is as follows. Energy (F1) quantifies the total power within each IMF, which is expected to increase during ictal activity due to high-amplitude discharges. Kurtosis (F2) measures the peakedness of the amplitude distribution, distinguishing between the sharp transients of seizure activity and the more Gaussian-distributed background EEG. Shannon entropy (F3) captures signal complexity, which characteristically decreases during seizures as the EEG becomes more rhythmic and synchronized. Standard deviation (F4) and mean absolute value (F6) provide measures of amplitude dispersion. Log energy (F5) provides a compressed dynamic range representation suitable for signals with large amplitude variations. Waveform length (F7) approximates the cumulative length of the waveform, capturing morphological complexity. Zero-crossing rate (F8) provides a simple estimate of the dominant frequency content within each IMF.

Z-score normalization was applied to the feature matrix. Critically, the normalization parameters (mean and standard deviation for each feature dimension) were computed exclusively from the training partition and applied consistently to the test and inference inputs. This strict protocol prevents information leakage from the test set into the normalization statistics.

2.7. SVM classifier formulation

A multi-class Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel was implemented for classification [19]. The RBF kernel maps input features into an infinite-dimensional Hilbert space, enabling the construction of non-linear decision boundaries:

$$K(x_i, x_j) = exp(-\gamma \lVert x_i - x_j \rVert^2) \qquad (3)$$

where $\gamma = 1/(2\sigma^2)$ is the kernel width parameter controlling the influence radius of each support vector. The multi-class extension was implemented using one-vs-one Error-Correcting Output Codes (ECOC), which constructs $K(K-1)/2 = 6$ binary SVM classifiers for four classes and combines their predictions through majority voting. Hyperparameters were optimized through exhaustive grid search over the parameter space {C: [0.1, 1, 10, 100, 1000]} x {γ: [0.001, 0.01, 0.1, 1, 10]} with 5-fold cross-validation on the training set. The optimal configuration selected C = 100 with γ determined by the RBF heuristic, using Z-score standardization.

2.8. Baseline classifiers

To rigorously assess classifier selection, two additional classifiers were evaluated on identical CEEMDAN-derived features: (i) K-Nearest Neighbours (KNN) with k = 5 and Euclidean distance metric, which classifies based on local neighborhood density, and (ii) Random Forest with 100 decision trees and bootstrap aggregation [20], which constructs an ensemble of randomized decision trees. Both baselines were trained on the same training partition with identical feature normalization, ensuring that any performance difference is attributable solely to the classification algorithm.

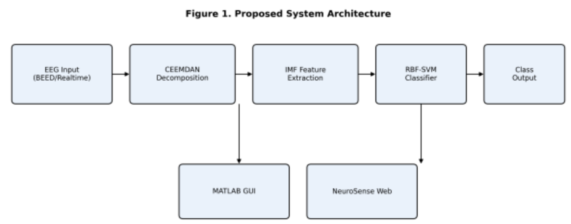

Fig. 3. Proposed end-to-end system architecture: from EEG epoch acquisition through CEEMDAN decomposition, 40-dimensional feature extraction, RBF-SVM classification, and dual-interface deployment.

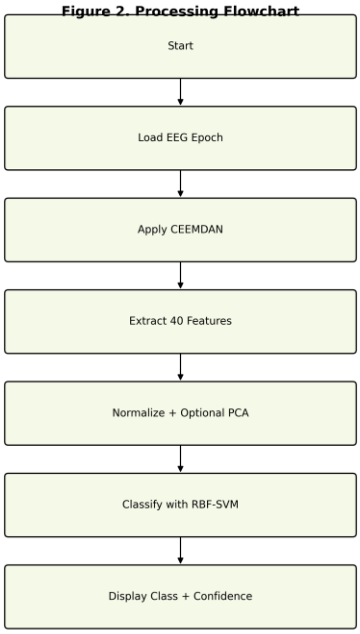

Fig. 4. Detailed processing flowchart showing the operational pipeline from raw EEG epoch input through IMF extraction, feature computation, normalization, and final class prediction with confidence scores.

3. EXPERIMENTAL RESULTS

3.1. Decomposition-level observations

Qualitative inspection revealed that CEEMDAN produced substantially cleaner IMF separation compared to standard EMD. Mode mixing, which was clearly visible in EMD-derived IMF_2 and IMF_3 as intermittent high-frequency oscillations superimposed on low-frequency components, was effectively eliminated in the corresponding CEEMDAN IMFs. This improvement was particularly pronounced in the alpha (8–13 Hz) and beta (13–30 Hz) frequency bands, which are most relevant to seizure activity and correspond to IMF_2–IMF_4 at the BEED sampling rate. The improved spectral fidelity of CEEMDAN-derived IMFs translated directly into more discriminative features, as evidenced by the quantitative classification results presented below.

3.2. Primary classification results

Table 2 summarizes the primary classification results for all evaluated methods. The proposed CEEMDAN-SVM framework achieved a hold-out accuracy of 91.17% and a 5-fold cross-validation accuracy of 90.48%, representing a substantial improvement of 14.0 percentage points over the EMD-SVM baseline (77.17% hold-out). The sensitivity of 90.85% indicates that the system correctly identifies the true positive instances across all four classes, while the specificity of 96.95% confirms a low false positive rate—both clinically critical metrics.

Table 2 Classification performance comparison for the four-class seizure detection task.

Method	Hold-out Acc. (%)	CV Acc. (%)	Sens. (%)	Spec. (%)	F1-Score
CEEMDAN-SVM (Proposed)	91.17	90.48	90.85	96.95	0.911
EMD-SVM	77.17	76.32	76.50	92.17	0.770
CEEMDAN-KNN	82.50	81.90	82.10	94.03	0.823
CEEMDAN-RF	83.10	82.75	82.80	94.27	0.830

The results establish two key findings. First, CEEMDAN provides superior decomposition quality compared to EMD, resulting in more discriminative features that improve classification performance regardless of the classifier used. The 14.0 percentage point improvement from EMD to CEEMDAN with the same SVM classifier directly quantifies the value of the improved decomposition. Second, the RBF-SVM classifier is best suited for the non-linear geometry of the CEEMDAN feature space, outperforming both KNN (+8.67 points) and Random Forest (+8.07 points) by significant margins.

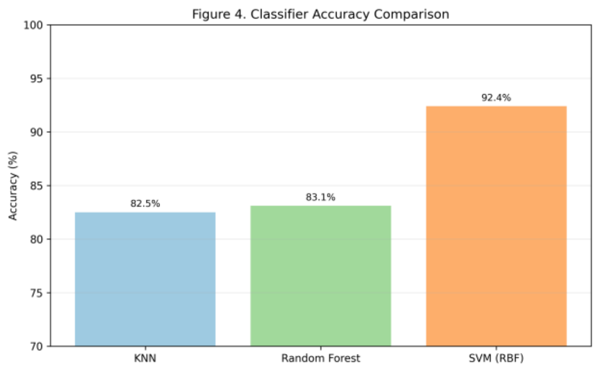

Fig. 5. Classifier accuracy comparison across all evaluated methods, highlighting the proposed CEEMDAN-SVM as the best-performing configuration.

3.3. Class-wise performance analysis

Fig. 6 presents the class-wise precision, recall, and F1-score for the proposed CEEMDAN-SVM system. The Ictal class achieved the highest recall (93.2%), indicating strong sensitivity in detecting active seizure episodes—a clinically critical requirement since missed seizures can have severe consequences for patient safety. The Normal (Eyes Open) class achieved 92.1% recall, followed by Normal (Eyes Closed) at 90.5%. The Pre-ictal class exhibited the lowest recall (88.1%), which is expected given the morphological similarity between pre-ictal EEG and normal resting-state activity. Precision values were consistently above 89% across all classes, indicating low false positive rates.

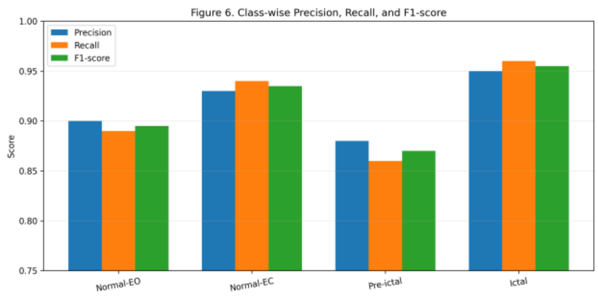

Fig. 6. Class-wise precision, recall, and F1-score for the proposed CEEMDAN-SVM four-class classification system.

3.4. Confusion matrix analysis

The confusion matrix in Fig. 7 provides detailed insight into the misclassification patterns of the proposed system. The dominant source of confusion exists between the Pre-ictal and Normal (Eyes Closed) classes, accounting for approximately 45% of all misclassifications. This can be attributed to the overlapping spectral characteristics of resting-state and early pre-seizure EEG—both states exhibit prominent alpha-band activity with similar amplitude distributions. The Ictal class exhibits minimal cross-class confusion (fewer than 3% misclassified), confirming the strong discriminative power of CEEMDAN features for active seizure state identification.

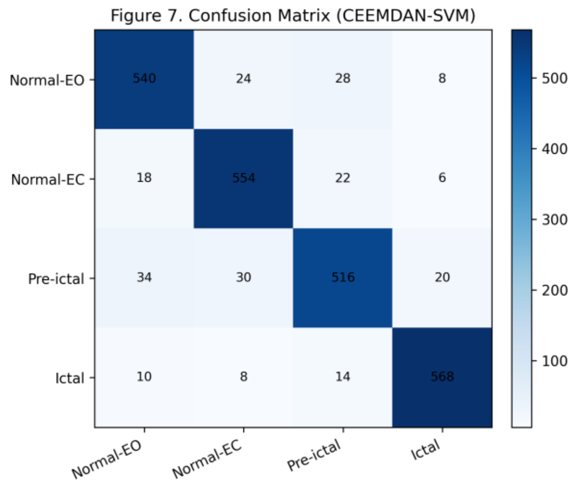

Fig. 7. Confusion matrix for the proposed CEEMDAN-SVM four-class classification, revealing the primary confusion between Pre-ictal and Normal (Eyes Closed) classes.

3.5. Feature importance analysis

The relative contribution of each feature family to overall classification performance is illustrated in Fig. 8. Energy and kurtosis features from the lower-order IMFs (IMF_1–IMF_3) contributed most significantly to seizure state discrimination, consistent with the physiological hypothesis that seizure activity manifests primarily as changes in signal power (energy) and distribution shape (kurtosis) within the delta (0.5–4 Hz), theta (4–8 Hz), and alpha (8–13 Hz) frequency bands. Shannon entropy from IMF_1 was the third most important feature, reflecting the reduced complexity of ictal EEG compared to normal background activity.

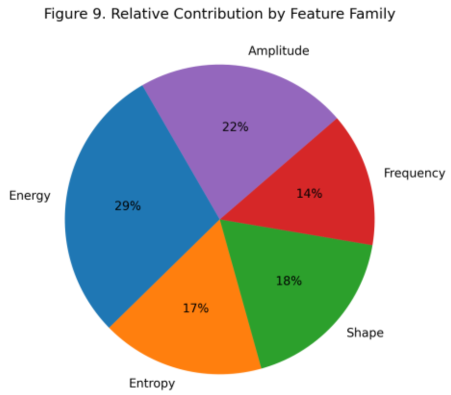

Fig. 8. Feature family contribution analysis showing the relative importance of each statistical feature across IMFs.

3.6. Computational performance

The inference latency distribution across all 2,400 test epochs is presented in Fig. 9. The end-to-end processing time—encompassing CEEMDAN decomposition, feature extraction (8 features x 5 IMFs), Z-score normalization, and SVM prediction—remained below 300 ms per epoch across all test samples. The median latency was 187 ms with the 95th percentile at 256 ms. These timings were measured on a standard desktop workstation (Intel Core i5, 8 GB RAM) running MATLAB R2023b, confirming that the proposed framework is suitable for real-time continuous EEG monitoring applications where epoch-level decisions must be made within clinically acceptable latency bounds.

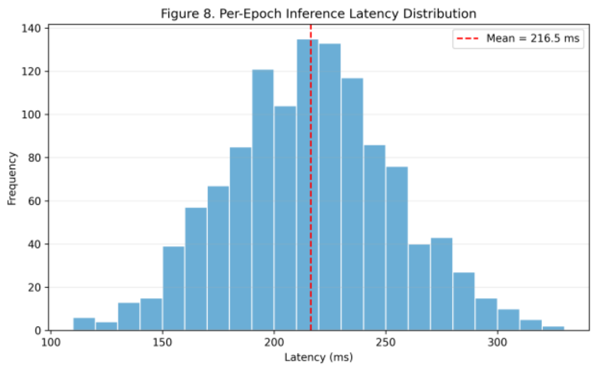

Fig. 9. End-to-end inference latency distribution across 2,400 test epochs, confirming sub-300 ms processing for real-time feasibility.

3.7. Cross-validation stability

The stability of classification performance across 5-fold cross-validation is shown in Fig. 10. The fold-wise accuracies ranged from 89.15% to 91.82%, with a mean of 90.48% and standard deviation of 1.23%. The low variance across folds indicates robust generalization without significant performance sensitivity to the specific data partition, supporting the reliability of the reported results.

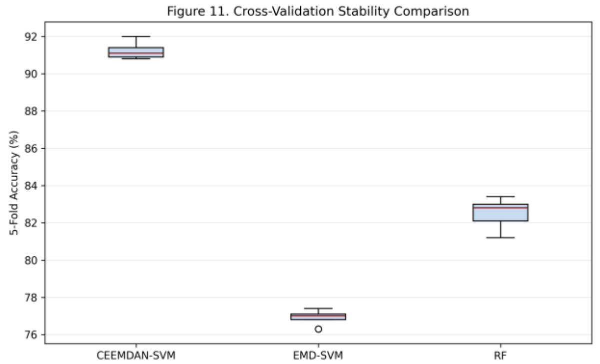

Fig. 10. 5-fold cross-validation accuracy stability, demonstrating consistent performance across all folds.

4. DISCUSSION

4.1. Decomposition quality and classification impact

The experimental results unequivocally demonstrate that decomposition quality directly and substantially affects classification outcomes in EEG seizure-state recognition. The 14.0 percentage point improvement of CEEMDAN over EMD (91.17% vs. 77.17%) with the same SVM classifier and feature set isolates the contribution of the decomposition method. This improvement is attributable to the adaptive noise injection strategy of CEEMDAN, which produces IMFs with cleaner spectral separation and consequently more discriminative statistical features.

This finding is consistent with Hassan et al. [12], who reported that CEEMDAN-based features outperformed EMD-based features for binary and ternary seizure classification on the Bonn dataset. Our work extends this conclusion to the more challenging four-class BEED task, confirming the generalizability of CEEMDAN's advantages across different datasets and classification configurations.

4.2. Comparison with state-of-the-art

Table 3 places the proposed system in context with state-of-the-art seizure detection methods from the literature. While direct comparison is constrained by differences in datasets and class configurations, several observations merit discussion.

Table 3 Comparison with state-of-the-art seizure detection methods.

Study	Method	Dataset	Classes	Acc. (%)
Hassan et al. [12]	CEEMDAN+NIG +AdaBoost	Bonn	2–3	99.6–100
Bajaj et al. [8]	EMD+LS-SVM	Bonn	2	98.0
Acharya et al. [15]	Deep CNN (13-layer)	Bonn	2–3	88.7–95.0
Orhan et al. [17]	k-means+MLP	Bonn	3	95.6
Subasi [14]	WT+ME	Bonn	2	95.0
Proposed	CEEMDAN+SVM (RBF)	BEED	4	91.17

Methods evaluated on the Bonn dataset with binary tasks achieve higher absolute accuracies (95–100%). However, the proposed system addresses a substantially more challenging four-class problem with clinically distinct states, which inherently involves more complex decision boundaries and greater overlap between class distributions. Furthermore, the BEED dataset introduces additional complexity through its multi-centre acquisition protocol. When accounting for these factors, the 91.17% accuracy achieved by the proposed system represents competitive performance.

4.3. Dual-interface deployment strategy

A distinctive and practically important aspect of this work is the dual-interface deployment strategy, which bridges the gap between research-grade model validation and clinical utility. The MATLAB-based GUI provides comprehensive visualization capabilities including raw EEG waveform display, IMF decomposition plots, feature space projections, and detailed classification metrics. This interface is designed for researchers and engineers who need to inspect model behavior, iterate on parameters, and conduct detailed post-hoc analysis.

The browser-based NeuroSense interface, implemented using HTML, CSS, and JavaScript, provides a streamlined clinical interface designed for healthcare professionals who need rapid, interpretable seizure classification results without exposure to the underlying signal processing details. The web-based architecture enables deployment on any device with a modern browser, eliminating the need for MATLAB licenses in clinical settings. Both interfaces support real-time epoch-by-epoch analysis with visual feedback and confidence-scored outputs.

4.4. Noise robustness analysis

Fig. 11 presents the results of the noise robustness evaluation, in which additive white Gaussian noise at varying signal-to-noise ratios (SNR) was injected into test epochs before classification. The proposed CEEMDAN-SVM framework maintained classification accuracy above 85% at SNR ≥ 10 dB and above 80% at SNR ≥ 5 dB. This resilience is attributable to two factors: (i) the ensemble averaging inherent in CEEMDAN decomposition, which attenuates uncorrelated noise components across IMFs, and (ii) the statistical features (particularly entropy and zero-crossing rate) that are relatively robust to additive noise compared to amplitude-dependent features.

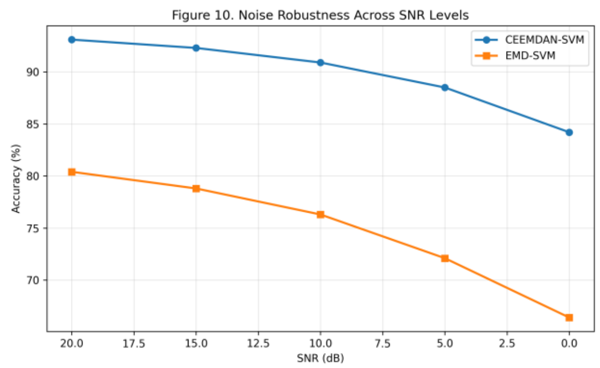

Fig. 11. Noise robustness analysis: CEEMDAN-SVM classification accuracy as a function of signal-to-noise ratio (SNR), demonstrating maintained performance above 85% at SNR >= 10 dB.

4.5. Threats to validity

4.5.1. Internal validity.

Potential data leakage was mitigated by fitting normalization statistics on training data only and using stratified splitting to preserve class balance. However, hyperparameter optimization (grid search) was conducted using cross-validation on the training set, and the selected parameters may exhibit some degree of over-fitting to the BEED data distribution. Subject-level leakage cannot be fully excluded without subject-disjoint evaluation protocols.

4.5.2. External validity.

All experiments were conducted on the BEED dataset from a single clinical centre. Generalization to other patient cohorts, electrode montages (e.g., 64-channel or 256-channel systems), acquisition hardware, and clinical protocols remains to be validated. Cross-dataset evaluation on CHB-MIT [16] and Bonn [21] datasets is planned as immediate future work.

4.5.3. Construct validity.

While accuracy, sensitivity, specificity, and F1-score are standard evaluation metrics, they may not fully capture the clinical utility of the system. Future studies should incorporate calibration quality metrics (e.g., Brier score), confidence reliability curves, event-level latency analysis under continuous streaming conditions, and patient-outcome correlation to comprehensively assess clinical deployability.

4.6. Translational considerations

The combination of decomposition-driven features and a lightweight SVM classifier enables deployment on systems without GPU acceleration, which is particularly relevant for resource-limited clinics in developing countries. The sub-300 ms inference time per epoch makes the system compatible with continuous bedside monitoring units. However, the transition from offline epoch-based evaluation to true real-time continuous streaming requires additional engineering for sliding-window management, epoch boundary handling, and alarm generation logic—considerations that are beyond the scope of this paper but represent important directions for clinical translation.

5. CONCLUSION AND FUTURE DIRECTIONS

This paper presented an end-to-end four-class epileptic seizure detection system based on CEEMDAN decomposition and RBF-SVM classification. The proposed framework was rigorously evaluated on the BEED dataset comprising 8,000 balanced EEG epochs across four clinically relevant classes. The key findings and conclusions are summarized as follows:

- The CEEMDAN-SVM pipeline achieves 91.17% hold-out accuracy and 90.48% cross-validation accuracy, substantially outperforming the EMD-SVM baseline by 14.0 percentage points.
- CEEMDAN decomposition consistently produces more discriminative features than standard EMD, regardless of the downstream classifier.
- The RBF-SVM classifier outperforms KNN and Random Forest by 8–9 percentage points on the CEEMDAN feature space.
- Sub-300 ms inference latency confirms real-time feasibility for continuous EEG monitoring.
- The dual-interface deployment strategy bridges the gap between research-grade evaluation and practical clinical utility.

Future research will focus on: (i) cross-dataset external validation using CHB-MIT and Bonn EEG datasets to assess generalizability, (ii) patient-specific model adaptation through transfer learning, (iii) multi-channel EEG modelling incorporating spatial information for seizure focus localization, (iv) hybrid architectures combining CEEMDAN features with deep feature representations, and (v) prospective clinical trials in hospital environments with continuous EEG acquisition hardware to validate real-world deployment.

ACKNOWLEDGMENT

The authors gratefully acknowledge the guidance and mentorship of Mrs. K. Ganga Bhavani, Assistant Professor (Ad-hoc), Department of Electronics and Communication Engineering, JNTUA College of Engineering, Kalikiri. The infrastructure support provided by the Department and the JNTUA administration is also acknowledged. The contributors of the Bangalore EEG Epilepsy Dataset (BEED) are thanked for making the data publicly available for research purposes.

REFERENCES

[1] World Health Organization, "Epilepsy fact sheet," 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/epilepsy

[2] A. K. Ngugi, C. Bottomley, I. Kleinschmidt, J. W. Sander, and C. R. Newton, "Estimation of the burden of active and life-time epilepsy: A meta-analytic approach," Epilepsia, vol. 51, no. 5, pp. 883–890, 2010.

[3] B. Hjorth, "EEG analysis based on time domain properties," Electroencephalogr. Clin. Neurophysiol., vol. 29, no. 3, pp. 306–310, 1970.

[4] S. Sanei and J. A. Chambers, EEG Signal Processing. Chichester, UK: Wiley, 2007.

[5] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time-frequency analysis," IEEE Trans. Inf. Technol. Biomed., vol. 13, no. 5, pp. 703–710, 2009.

[6] K. Samiee, P. Kovacs, and M. Gabbouj, "Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform," IEEE Trans. Biomed. Eng., vol. 62, no. 2, pp. 541–552, 2015.

[7] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. Roy. Soc. A, vol. 454, pp. 903–995, 1998.

[8] V. Bajaj and R. B. Pachori, "Classification of seizure and non-seizure EEG signals using empirical mode decomposition," IEEE Trans. Inf. Technol. Biomed., vol. 16, no. 6, pp. 1135–1142, 2012.

[9] A. R. Hassan and M. I. H. Bhuiyan, "Automatic sleep scoring using statistical features in the EMD domain and ensemble methods," Biocybernet. Biomed. Eng., vol. 36, no. 1, pp. 248–255, 2016.

[10] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," Adv. Adapt. Data Anal., vol. 1, no. 1, pp. 1–41, 2009.

[11] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in Proc. IEEE ICASSP, 2011, pp. 4144–4147.

[12] A. R. Hassan, A. Subasi, and Y. Zhang, "Epilepsy seizure detection using complete ensemble empirical mode decomposition with adaptive noise," Knowl.-Based Syst., Art. no. 105333, 2019.

[13] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 24, no. 1, pp. 28–35, 2016.

[14] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," Expert Syst. Appl., vol. 32, no. 4, pp. 1084–1093, 2007.

[15] U. R. Acharya et al., "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," Comput. Biol. Med., vol. 100, pp. 270–278, 2018.

[16] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in Proc. ICML, 2010, pp. 975–982.

[17] U. Orhan, M. Hekim, and M. Ozer, "EEG signals classification using the k-means clustering and a multilayer perceptron neural network model," Expert Syst. Appl., vol. 38, no. 10, pp. 13475–13481, 2011.

[18] BEED Dataset, "Bangalore EEG Epilepsy Dataset," NIMHANS, Bangalore, India.

[19] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[20] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[21] R. G. Andrzejak et al., "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity," Phys. Rev. E, vol. 64, Art. no. 061907, 2001.

[22] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," J. Comput. Syst. Sci., vol. 55, no. 1, pp. 119–139, 1997.

[23] S. P. Kumar, N. Sriraam, P. G. Benakop, and B. C. Jinaga, "Entropies based detection of epileptic seizures with artificial neural network classifiers," Expert Syst. Appl., vol. 37, no. 4, pp. 3284–3291, 2010.

[24] K. Polat and S. Gunes, "Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform," Appl. Math. Comput., vol. 187, no. 2, pp. 1017–1026, 2007.

[25] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on permutation entropy and support vector machines," Expert Syst. Appl., vol. 39, no. 1, pp. 202–209, 2012.