SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I8P130

Volume 12 Issue 8, 352-360, August 2025

Original Article

Accurate Epileptic Seizure Detection from EEG Using Feature Fusion and MI-Enhanced XGBoost Classifier

Mamatha G N[1], Hariprasad S A[2]

[1,2]Department of Electronics and Communication Engineering, Faculty of Engineering and Technology, Jain University, Bengaluru, Karnataka, India.

[1]Corresponding Author : gn.mamatha@jainuniversity.ac.in

Received: 16 June 2025 Revised: 17 July 2025 Accepted: 18 August 2025 Published: 30 August 2025

Abstract - *The accuracy of Electroencephalogram (EEG) signal-based epileptic seizure identification is often compromised by poor feature selection and duplicate data. This research proposes a method that combines early feature fusion from many domains with Mutual Information (MI)-based feature selection to overcome these issues. Principal Component Analysis (PCA), Hilbert–Huang Transform (HHT), Reconstruction Independent Component Analysis (RICA), and Empirical Mode Decomposition (EMD) are used to extract features that capture time, frequency, and nonlinear information. The Extreme Gradient Boosting (XGBoost) algorithm is used to categorize the most relevant qualities once Mutual Information has been utilized to choose them. The suggested approach performs exceptionally well on all significant measures when using the Bonn EEG dataset. Its efficient design ensures both enhanced detection capability and suitability for real-time clinical use.*

Keywords - *Epilepsy, EEG, Feature fusion, Mutual information, XG-Boost, BONN Dataset.*

1. Introduction

Epileptic seizures, a neurological disorder that affects over 65 million individuals globally, must be identified early and accurately to be effectively treated and managed [1]. Electroencephalography (EEG) is still an essential technique for tracking brain activity and detecting seizures. However, manual analysis is challenging and error-prone because EEG signals are complicated and non-stationary. As a result, the demand for sophisticated and automated seizure detection techniques has grown. While a number of previous studies have investigated seizure detection using individual techniques, such as Hilbert-Huang Transform (HHT), Multivariate EMD with neural networks, Empirical Mode Decomposition (EMD) [2-4], or even XGBoost with single-domain features [5, 6], these methods frequently lack a comprehensive representation of the signal or are unable to optimize feature selection efficiently. Interestingly, none of these approaches have combined multi-domain features by using Mutual Information (MI)-based feature selection [7].

This study suggests a unique framework that uses PCA, HHT, EMD, and RICA to extract various features, followed by early fusion to retain inter-domain interactions, to solve the performance gap in EEG seizure detection. Then, only the most pertinent characteristics are kept for classification using XGBoost by applying Mutual Information (MI)-based feature selection. When tested on the BONN EEG dataset, the approach outperforms current state-of-the-art methods in terms of accuracy and shows great promise for enhancing automated seizure detection systems.

1.1. Paper Organization

Section 2 examines current methods for extracting features from EEGs based on epilepsy. Section 3 describes the suggested approach, which combines XGBoost classification, MI-based feature selection, four feature extraction techniques with early fusion, and the BONN dataset. While Section 5 summarizes the main conclusions and suggests future study topics, Section 4 displays and contrasts the experimental outcomes.

2. Relevant Brief Description

One important area of research aimed at facilitating early diagnosis and efficient monitoring is EEG-based epileptic seizure detection. Traditional analysis techniques are challenged by the complex, nonlinear, and non-stationary character of EEG signals, which leads to the creation of sophisticated frameworks for feature extraction, fusion, selection, and classification.

2.1. Methods for Feature Extraction

One of the first steps in seizure detection for the preprocessed signal is the feature extraction. Discrete Wavelet Transform and Welch's Power Spectral Density are two common frequency-domain methods for representing spectral energy patterns [8]. While nonlinear properties like fractal dimension, sample entropy, and permutation entropy reflect the inherent complexity of brain dynamics [10], time-domain statistical measures like mean, variance, skewness, and kurtosis aid in capturing transient signal characteristics [9]. Additionally, adaptive, multi-resolution analysis using Intrinsic Mode Functions (IMFs) is made possible by data-driven decompositions such as HHT and EMD [11-14].

2.2. Reducing Dimensionality and Choosing Features

High-dimensional feature spaces may result in overfitting and redundancy. To maintain variance or independence across components, dimensionality reduction techniques such as PCA and ICA are frequently employed [15]. The feature space is further refined using optimization-based techniques, including Pearson correlation analysis, Grasshopper Optimization Algorithm (GOA), and Particle Swarm Optimization (PSO) [16, 17].

Mutual Information (MI), which may capture both linear and nonlinear correlations between features and class labels, is a potent method for assessing feature significance, according to recent studies [8, 17, 18]. Even though MI can improve classifier performance and signal representation, it is currently used infrequently after multi-domain feature fusion.

2.3. Methods of Feature Fusion

To improve resilience and discriminative capacity, feature fusion algorithms combine complementary data from the temporal, frequency, and nonlinear domains. Studies like [5, 7] have shown that combining statistical and spectral data with XGBoost improves detection accuracy. Hybrid fusion in conjunction with MI-based feature selection lowers computational load and enhances generalizability, as shown by Subasi et al. [18, 19]. Like this, writers in [20-22] verified that early multi-domain feature fusion greatly improves classification performance when combined with efficient selection techniques. Many current methods, however, lack a structured pipeline for fusion and selection, which results in high computational demands, redundant features, and overfitting problems that are particularly significant in real-time applications.

2.4. Models of Classification

Following the retrieval and selection of features, classifiers such as an ensemble tree-based approach called XGBoost have demonstrated better performance in managing high-dimensional fused features and preventing overfitting in more recent times [6, 25, 26]. It is perfect for real-time or embedded seizure detection applications due to its stability and scalability.

Current Limitations in seizure detection methods are:

- Single-domain feature dependence
- Inadequate feature selection following fusion
- Use of Mutual Information (MI) to improve feature relevance is limited.

2.4.1. Proposed Methodology Suggests

Early feature extraction and fusion from several domains enhances the representation using:

- Principal Component Analysis (PCA)
- Hilbert-Huang Transform (HHT)
- Reconstructed Independent Component Analysis (RICA)
- Empirical Mode Decomposition (EMD)

2.4.2. Feature Selection

- Choosing Features: MI is used to choose the most informative features, which increases classification performance and decreases redundancy.

2.4.3. Classification

- The features' robustness is also evaluated and compared on the BONN dataset using a classification technique, namely XGBoost.

3. Methodology

To guarantee precise and dependable epileptic seizure detection, a methodical experimental framework was created by combining sophisticated signal processing techniques with strong machine learning methodologies. The main goals are reducing feature redundancy, improving classification performance, and identifying significant patterns in EEG signals.

The entire process is shown in Figure 1, which details the steps that must be followed to guarantee the best possible feature representation and increased detection accuracy: data preparation, feature extraction, feature selection, and classification.

The suggested seizure detection framework was assessed using the preprocessed Bonn EEG dataset. This dataset includes five 23.6-second single-channel EEG sets (A-E): Sets A and B (healthy people, eyes open and closed, respectively); Set E (ictal/seizure activity); and Sets C and D (interictal/seizure-free from epileptic patients). Utilizing this preprocessed artifact-free and validated dataset (presented in 27) guarantees the study's dependability.

After preprocessing, four complementary approaches encompassing time, frequency, and nonlinear properties are used in feature extraction.

- ✓ Adaptive time-frequency characteristics are extracted from non-stationary EEG signals using HHT and EMD.
- ✓ PCA preserves important variance components while reducing noise.
- ✓ Independent and sparse signal patterns are captured by RICA.

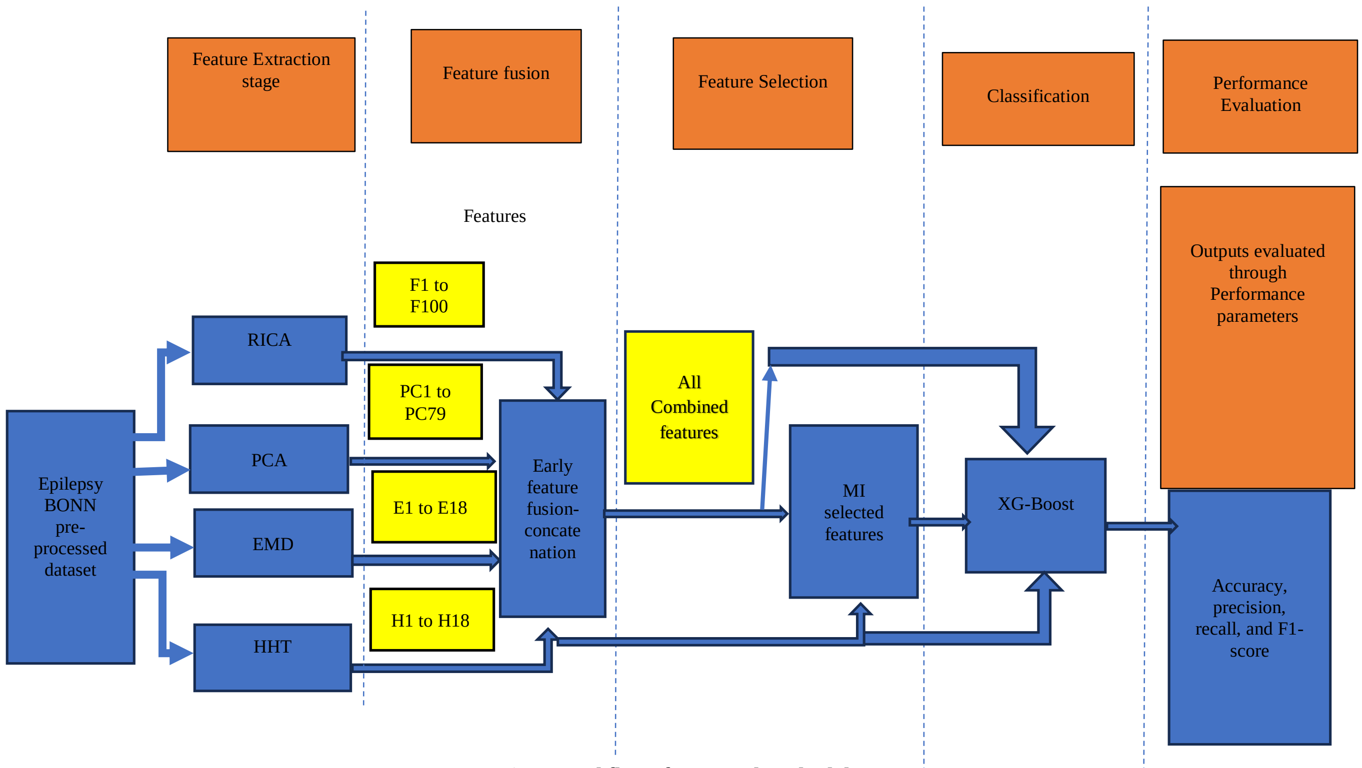

Fig. 1 Workflow of proposed methodology

Mutual Information (MI) is applied for feature selection, minimizing dimensionality and preventing overfitting by identifying the most pertinent characteristics. The chosen characteristics are then categorized using XGBoost, an effective, scalable, and quick method that works well with high-dimensional data. Model performance is examined using established classification techniques to ensure dependability. The "Results and Discussion" section will examine how each feature type affects seizure detection accuracy, while the following sections will describe these methods and how they operate. This systematic approach aims to create a reliable and efficient system that will serve as a solid basis for further studies in EEG-based seizure identification.

3.1. Methods of Feature Extraction and Early Fusion

Using early feature fusion, the suggested approach combines features from four sophisticated signal processing methods, PCA, HHT, EMD, and RICA. This method produces a more robust and discriminative feature representation by capturing complementary information from EEG signals throughout the time, frequency, and time-frequency domains [7-9].

3.1.1. Reconstructed Independent Component Analysis (RICA)

Given a multivariate EEG signal $X=[x_1,x_2,\ldots,x_T]^T \in R^{(T\times N)}$, where T is the number of time points and N is the number of channels (for the BONN dataset, N=1), RICA aims to find a de-mixing matrix $W \in R^{(K\times N)}$ such that the source signals $S=XW^T \in R^{(T\times K)}$ are statistically independent [27]. Here, K is the number of Independent Components (ICs). The reconstruction of the signal using a subset of P selected ICs (where P≤K) and a corresponding mixing matrix $A \in R^{(N\times P)}$ can be represented as in Equation (1):

$$\hat{x} = S_p A^T \quad (1)$$

Where S_P contains the P selected ICs. After that, extract statistical features f_{RICA} from each reconstructed component X, which is the i^{th} column of $\hat{x}$, such as mean, standard deviation, skewness and kurtosis, etc. The feature set from RICA is F_{RICA} = [Mean ($\hat{x}i$)…… Kurtosis($x^{>}$)].

3.1.2. Hilbert-Huang Transform (HHT)

EMD decomposes the signal x(t) into a sum of Intrinsic Mode Functions (IMFs) $c_i(t)$ and is mentioned in Equation (2):

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \quad (2)$$

Let the residue be represented by $r_n(t)$. Two requirements must be met by each Intrinsic Mode Function (IMF) $c_i(t)$:

1. There must be no more than one difference between the number of extrema and zero crossings, or they must be equal.
2. The mean value of the envelopes that are defined by the local minima and maxima must always be zero.

The Hilbert Transform $H\{c_i(t)\}$ of an IMF $c_i(t)$ is given in Equation (3):

$$H\{c_i(t)\} = \frac{1}{\Pi} P \cdot v \int_{-\infty}^{\infty} \frac{c_i(T)}{t-T} \cdot dT \quad (3)$$

Where P.V. denotes the Cauchy principal value.

3.1.3. Empirical Mode Decomposition (EMD)

As described above, EMD yields a set of IMFs $C_i(t)$. Extracting statistical features f_{EMD} from the first few IMFs, such as energy $Ei = \sum_{t=1}^{T} |c_i(t)|^2$ And Shannon Entropy is the normalized energy at time t for the i^{-th} IMF. The feature set is F_{EMD}= [E_1, ..., S_M], where M is the number of considered IMFs represented in Equation (4).

$$S_i = -\sum_{t=1}^{T} P_{it} \, log(P_{it}) \quad \text{where } P_{it} = \frac{|C_i(t)|^2}{E_i} \tag{4}$$

3.1.4. Principal Component Analysis (PCA)

The goal of PCA is to find the Set of orthogonal principal components that best captures the variation in the EEG data matrix X. The covariance matrix C, which shows the variances and correlations between the EEG features, is mentioned in Equation (5). The following formula is used to determine the covariance matrix:

$$C = \frac{1}{T-1}(X - \bar{x})^T (x \quad - \bar{x}) \tag{5}$$

Where $\bar{x}$ The mean vector of the columns of X and T is the number of observations.

Through the identification of key patterns, PCA lowers the dimensionality of data. This is accomplished by calculating the covariance matrix's eigenvectors and eigenvalues. Equation (6) shows that the top Q eigenvectors, which correspond to the biggest eigenvalues, are chosen because they capture the greatest amount of variance. The data is subsequently transformed using these chosen eigenvectors.

$$Y = XV_Q \tag{6}$$

The selected eigenvectors found in V Q are then used to define this lower-dimensional space onto which the original data, X, is projected. The most important features of the data are retained in this projection, which produces a simpler representation [10, 18].

The key components of an EEG signal are captured by extracting essential features. Signal shape and intensity are summarized by statistical measures (mean, median, standard deviation, minimum, maximum, and energy). Signal strength, frequency content, and temporal change are reflected in the Hjorth parameters (activity, mobility, and complexity). The evolution of the signal is traced by cumulative features (cumulative mean, minimum, and maximum). Signal unpredictability is assessed using entropy measurements (Shannon, Rényi, Approximate, and Sample Entropy). Lastly, self-similarity and nonlinear complexity are quantified by fractal dimension features (Higuchi, Katz).

3.1.5. Early Feature Fusion

Concatenating the feature vectors derived from each technique is the first stage in feature fusion: F_{fused} = [FRICA, FHHT, FEMD, FPCA]. The product is a thorough feature vector incorporating data from various signal processing domains.

3.1.6. Feature Selection Based on Mutual Information (MI)

An information-theoretic metric called Mutual Information (MI) measures the statistical dependency between two random variables. The definition of the mutual information I(X, Y) for continuous variables X and Y is as follows [19, 28]:

$$I(X;Y) = \frac{\int\int ((pX,Y(x,y)log(pX,Y(x,y))}{(pX(x)pY(y)))dx.dy} \tag{7}$$

From Equation (7), where the combined probability density function of X and Y is represented by pXY(x,y), and the marginal distributions of X and Y are represented by pX(x) and pY(y).

MI is computed between each feature in the fused feature set and the class label (seizure or non-seizure). Because it exhibits a stronger correlation with the class name, a feature with a higher MI value is more crucial for differentiating between the two classes.

3.2. Classifier Stage

XGBoost is a reliable and appropriate classifier for classifying EEG signals, especially in seizure detection. As a result, the following section will offer a thorough explanation and performance evaluation. To fully assess the effectiveness of the suggested framework, a comparison analysis carried out during the feature extraction step across four different settings will be summarized in the "Results and Discussion" section.

Every scenario's performance metrics were evaluated and documented for the selected classifiers, such as Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbour (KNN) and XG-boost. The impact of early feature fusion and MI-based feature selection on the XGBoost model's overall performance will be covered in this section.

3.2.1. Extreme Gradient Boost Classifier

In this study, the high-performance XGBoost algorithm is used to classify optimal features that were chosen using Mutual Information (MI). An ensemble of decision trees is constructed successively using XGBoost, each of which fixes the mistakes of the one before it. This method's speed, scalability, and integrated regularization make it perfect for handling complicated EEG data. Equation (8) [29] illustrates how XGBoost effectively detects seizures by spotting nonlinear patterns and reducing classification errors, which reduces the possibility of overfitting.

$$y_i = \sum_{k=1}^{K} f_k(x_i) \tag{8}$$

3.2.2. XGBoost Algorithm Steps for EEG-Based Epilepsy Classification

Input Data Preparation

- Input: Preprocessed EEG signals (e.g., from the Bonn dataset).
- Features: Extracted from time, frequency, time-frequency, or nonlinear domains.
- Apply feature selection (e.g., MI) to retain only the most relevant features.

Data Splitting

- Create distinct training and testing sets from the dataset (usually 70–30 or 80–20).
- Optionally use k-fold cross-validation for better model generalization.

Initialize Base Learners

- Decision trees (usually Classification And Regression Trees, or CART) are used by XGBoost as base learners.
- Set initial prediction (often the mean log odds or class prior probability).

Train Trees Iteratively For each boosting round (iteration)

- Compute Gradient and Hessian:
  For each data point, calculate the gradient (1st derivative) and hessian (2nd derivative) of the loss function (usually log loss for classification).
- Construct Decision Tree:
  Build a tree that best splits the data based on the gradient and hessian, maximizing gain (reduction in loss).
- Regularization:
  Apply penalties to tree depth, leaf weights, and number of leaves to avoid overfitting (controlled by parameters like lambda, alpha, and max_depth).
- Update Predictions:
  Add the new tree's weighted predictions to the existing model.

Stopping Criteria

- Stop when the maximum number of trees (n_estimators) is reached, or if improvement in loss falls below a threshold (early stopping).

Model Output

- For classification: A probability score for each class is the end result.
- To determine whether a person is epileptic or not, apply a threshold (such as 0.5).

Evaluation

- Metrics such as accuracy, precision, recall, and F1-score are used to evaluate performance.

3.3. Performance Evaluation

Four important metrics, accuracy, precision, recall, and F1-score, were used to evaluate the classification models' effectiveness [30]. By statistically assessing the model's capacity to discriminate between seizure and non-seizure events, these metrics demonstrate the model's overall efficacy, robustness, and dependability in epilepsy detection.

3.3.1. Accuracy

By calculating the percentage of all predictions that are correctly classified across all classes, accuracy provides insight into the model's overall correctness.

$$\text{Accuracy} = \frac{\mathrm{T_p}+\mathrm{T_n}}{\mathrm{T_p}+\mathrm{T_n}+\mathrm{F_p}+\mathrm{F_n}} \tag{9}$$

Where:
True Positive (TP): The number of seizure episodes that the model properly classified as seizures.
True Negative (TN): To what extent were non-seizures accurately anticipated to be non-seizures?
False Positive (FP): The number of incidents incorrectly classified as seizures but not seizures.
False Negative (FN): How many actual seizure events did the model miss because it assumed they were non-seizures?

3.3.2. Recall (Sensitivity) (True Positive Rate)

The model's recall gauges how well it can detect positive samples, or seizure occurrences.

$$\text{Recall} = \frac{\mathrm{T_p}}{\mathrm{T_p}+\mathrm{F_n}} \tag{10}$$

3.3.3. Precision (Positive Predictive Value)

A measure of precision is the proportion of accurately recognized seizure occurrences among all events predicted to be seizures (i.e., true positives / (true positives + false positives)). It demonstrates how the model reduces false alarms by successfully differentiating seizures from non-seizures.

$$Precision = \frac{T_p}{T_p+F_p} \tag{11}$$

3.3.4. F1-Score

The F1-score is the harmonic mean of recall and precision. A higher F1-score indicates a well-balanced sensitivity and precision of the model.

$$F1 - score = 2 * \frac{P*R}{P+R} \tag{12}$$

Where: P= Precision, R = Recall

Because it guarantees that both recall and accuracy are considered in the evaluation, the F1-score is particularly helpful when the dataset is unbalanced.

3.3.5. Specificity

The model's ability to identify non-seizure events is measured by its specificity, also known as the true negative rate.

$$\text{Specificity} = TN / (TN + FP) \tag{13}$$

4. Results and Discussion

The four previously stated feature extraction and selection techniques were applied to the EEG data after hybrid EMD-ICA preprocessing to assess the effectiveness of the suggested seizure detection system. Together, these complementary methods produced 215 distinguishing characteristics. MI was used to rank each feature's importance in relation to seizure classification to improve model performance and minimize feature redundancy. A refined set of 136 highly relevant characteristics was obtained by eliminating 79 less informative features and keeping those whose MI scores were higher than the 5% threshold.

4.1. Experimental Results and Performance Assessment

A well-known machine learning classifier called XGBoost was used to evaluate the effectiveness of the proposed approach. Its performance was assessed using the conventional metrics of F1-score, accuracy, precision, and recall under four different scenarios.

1. Individual Feature Sets (HHT) without MI– Features obtained from the HHT extraction technique without applying Mutual Information.
2. Individual Feature Sets (HHT) with MI – The same features were refined using MI for relevance and redundancy reduction.
3. Combined Feature Set without MI – Early fused features from all extraction techniques (PCA, HHT, RICA, EMD) without selection.
4. Combined Feature Set with MI – Fused multi-domain features followed by MI-based selection.

4.1.1. HHT Feature Classification (Prior to MI)

Figure 2 displays the 401 × 18 feature matrix created for the features taken from the selected epileptic EEG recordings using the HHT technique. The selected classifier performance results are summarized in Table 1 following their training with these extracted attributes.

	A	B	C	O	P	Q	R
1	HHT features						
2	H_Mean	H_Median	H_Std_D	H_SampEr	H_energy	H_HFD	H_KFD
3	0.136647	0.109785	0.109505	0.208015	76.93179	0.144028	0.125129
4	0.549094	0.434484	0.45824	0.199046	1138.21	0.145465	0.125917
5	0.132394	0.106967	0.107828	0.164979	67.57759	0.117558	0.100107
6	0.267613	0.21729	0.218204	0.165121	290.9121	0.117056	0.100335
7	0.815139	0.672876	0.607676	0.238244	2741.367	0.147287	0.126609
8	0.192495	0.160045	0.136598	0.246963	146.6575	0.144502	0.125224
9	0.123841	0.100835	0.096511	0.175864	58.36601	0.121156	0.1001
10	0.121885	0.101681	0.087593	0.223106	61.7173	0.133117	0.111217
11	0.189602	0.152018	0.154232	0.243301	152.6523	0.165352	0.143113

Fig. 2 HHT features Set Prior to MI-based Feature Selection.

Table 1. Classification using HHT feature set (Before MI selection)

1	Accuracy	Recall	Precision	F1_score
SVM	0.8925	0.9012	0.9142	0.9024
DT	0.9650	0.9460	0.9722	0.9799
KNN	0.3175	0.4142	0.3575	0.2087
XGB	0.9750	0.9837	0.9710	0.9765

As shown in Table 1, with an accuracy of 97.50% and an F1-score of 97.65%, the model demonstrated performance, suggesting that HHT can also effectively extract significant patterns from EEG signals to classify seizures.

4.1.2. Classification Using HHT Features After MI Selection

Before being utilized for classification, the HHT features were initially processed using MI to identify the most pertinent aspects, as represented in Figure 3. These features have been employed with the same XG-Boost classifier. Table 2 displays the performance outcomes following MI-based feature selection.

	A	B	C	D	E	F	G	H	I	J
1	HHT best features									
2	H_ShanEn	H_SampEr	H_energy	H_KFD	H_Cmean	H_apprEn	H_Median	H_Activity	H_Mobilit	H_Std_D
3	256.0388	0.208015	76.93179	0.125129	241.748	0.280666	0.109785	0.05392	0.136334	0.109505
4	-1793.57	0.199046	1138.21	0.125917	907.2837	0.268917	0.434484	0.795565	0.139144	0.45824
5	216.4944	0.164979	67.57759	0.100107	220.0735	0.221007	0.106967	0.058992	0.119161	0.107828
6	-283.403	0.165121	290.9121	0.100335	437.8415	0.230232	0.21729	0.286098	0.121782	0.218204
7	-4429.26	0.238244	2741.367	0.126609	1430.891	0.308484	0.672876	1.781393	0.159123	0.607676
8	103.5653	0.246963	146.6575	0.125224	331.8622	0.330988	0.160045	0.091869	0.146378	0.136598
9	254.4401	0.175864	58.36601	0.1001	203.5445	0.238729	0.100835	0.051397	0.130142	0.096511
10	220.6276	0.223106	61.7173	0.111217	210.8791	0.296912	0.101681	0.045752	0.147591	0.087593
11	154.6462	0.243301	152.6523	0.143113	329.0819	0.328662	0.152018	0.094449	0.15581	0.154232

Fig. 3 HHT features set after MI-based feature selection

Table 2. Classification using HHT feature set (After MI-based selection)

2.	Accuracy	Recall	Precision	F1_score
SVM	0.9025	0.9120	0.9142	0.9021
DT	0.9550	0.9420	0.9622	0.9699
KNN	0.3375	0.4242	0.3675	0.2187
XGB	0.9805	0.9916	0.9805	0.9883

The effects of using MI-based feature selection on the same HHT feature set are shown in Table 2. This resulted in a discernible increase, increasing the F1-score to 98.83% and the accuracy to 98.05%. This illustrates how MI improves the classifier's overall performance by eliminating redundancy and keeping the most important features. The confusion matrix for the classifier trained solely using HHT characteristics is displayed in Figure 4. According to the matrix, there was one misclassification in which a healthy EEG was mistakenly predicted to be interictal, even though the seizure and interictal classes were accurately identified.

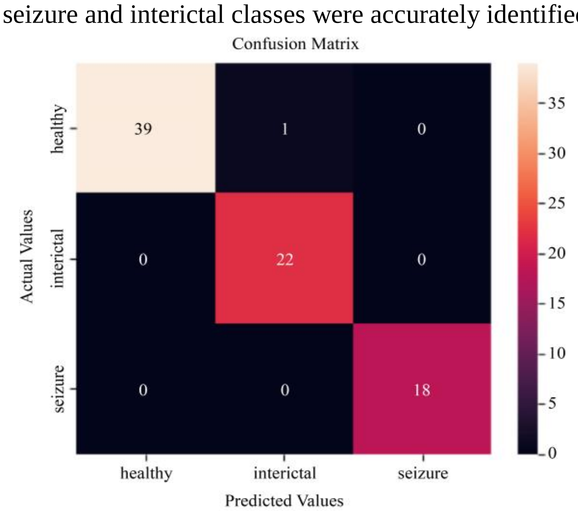

Fig. 4 Confusion matrix for the classifier trained on the HHT feature set for the XG boost classifier

4.1.3. Classification Analysis Using Combined Features: Early Feature Fusion without applying MI

Early feature fusion was used to merge features from the four chosen feature extraction methods: RICA, PCA, EMD and HHT. Figure 5 represents the resulting subset of this extensive feature matrix, which has 401 × 215 features. The associated performance outcomes were noted and shown in Table 3 following classification using these fused features.

	A	CV	CW	FW	FX	GO	GP	HG
1	RICA features		PCA features		HHT features		EMD features	
2	F1	F100	PC1	PC79	H_Mean	H_KFD	E_Mean	E_KFD
3	0.053303	-0.04405	7.97375	-0.19521	0.136647	0.125129	2.24E-05	1.009384
4	0.031674	-0.10594	64.53699	0.402626	0.549094	0.125917	0.000899	1.038863
5	0.014676	0.029994	-13.3657	0.753841	0.132394	0.100107	0.000161	1.011606
6	0.023426	0.021493	-94.1092	0.424399	0.267613	0.100335	0.003757	1.024794
7	0.00694	-0.00812	173.5653	0.374099	0.815139	0.126609	0.009128	1.055085
8	-0.07787	-0.01303	-51.2432	-0.05235	0.192495	0.125224	0.000784	1.014751
9	-0.01676	0.091984	-21.9633	0.050154	0.123841	0.1001	0.005183	1.010499
10	-0.14829	0.015968	-26.5228	2.852922	0.121885	0.111217	0.001364	1.009617
11	0.089914	0.019335	42.96941	0.366618	0.189602	0.143113	0.001756	1.012327
12	0.00537	0.069641	-72.6466	0.204436	0.36343	0.143545	0.003373	1.025872

Fig. 5 Combined feature set prior to MI -based feature selection

Table 3. Classification using the combined feature set (Prior to MI-based selection)

3.	Accuracy	Recall	Precision	F1_score
SVM	0.9125	0.9220	0.9242	0.9021
DT	0.9550	0.9420	0.9522	0.9499
KNN	0.3975	0.4342	0.3775	0.2387
XGB	0.9875	0.9761	0.9815	0.9686

The performance with this early feature fusion without MI-based selection is shown in Table 3. With an accuracy of 98.75% and an F1-score of 96.86%, this setup produced better results. This result emphasizes the value of combining characteristics from different domains since complementary information from different extraction techniques produces a more robust data representation.

4.1.4. Analysis of Combined Features for Classification: Early Feature Fusion with MI selection

Figure 6 displays the most pertinent features that are highly correlated with seizure activity following MI-based selection. The equivalent performance outcomes were then recorded and shown in Table 4 after these MI-selected features were put into the XG-Boost classifiers.

	A	U	V	CV	CW	DN	DO	EF
1	Combined best features							
2	RICA FEATURES		PCA FEATURES		HHT FEATURES		EMD FEATURES	
3	F3	F93	PC1	PC79	H_Mean	H_KFD	E_Mean	E_KFD
4	-0.04681	0.033301	7.97375	-0.19521	0.136647	0.125129	2.24E-05	1.009384
5	0.004043	-0.42812	64.53699	0.402626	0.549094	0.125917	0.000899	1.038863
6	-0.02652	0.136628	-13.3657	0.753841	0.132394	0.100107	0.000161	1.011606
7	0.028129	-0.12026	-94.1092	0.424399	0.267613	0.100335	0.003757	1.024794
8	-0.02656	-0.22336	173.5653	0.374099	0.815139	0.126609	0.009128	1.055085
9	0.143961	0.137034	-51.2432	-0.05235	0.192495	0.125224	0.000784	1.014751
10	-0.07766	-0.02439	-21.9633	0.050154	0.123841	0.1001	0.005183	1.010499
11	-0.24218	0.03207	-26.5228	2.852922	0.121885	0.111217	0.001364	1.009617
12	0.018723	0.017075	42.96941	0.366618	0.189602	0.143113	0.001756	1.012327

Fig. 6 Feature set after MI-based dimensionality reduction

Table 4. Classification performance using the combined feature set (After MI-based selection)

4.	Accuracy	Recall	Precision	F1_score
SVM	0. 9425	0.9132	0.9242	0.9024
DT	0.9650	0.9460	0.9722	0.9699
KNN	0.3975	0.4542	0.3575	0.2487
XGB	0.9985	0.9861	0.9914	0.9886

The classifier trained on early fused features demonstrated perfect seizure EEG classification and an improved balance across classes after MI-based feature selection, with minimal healthy/interictal misclassifications (Figure 7).

This last scenario produced nearly flawless XGBoost classifier metrics (Accuracy, Recall, Precision, F1-score = 99.85%) by combining early feature fusion with MI-based selection (Table 4). Although quite successful, it is important to recognize that results vary depending on the dataset. More testing on a variety of datasets is recommended to guarantee generalizability and reduce any biases or overfitting.

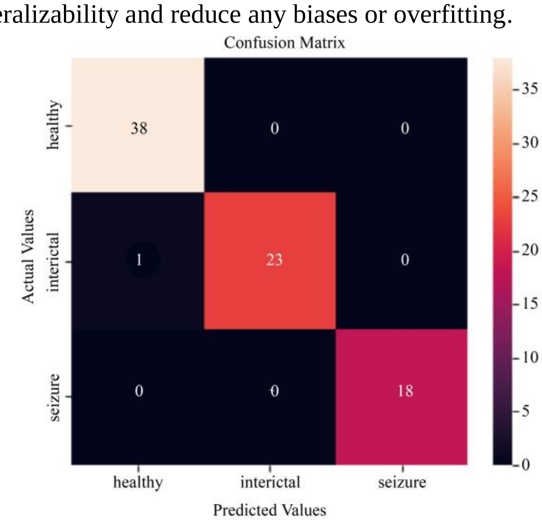

Fig. 7 Confusion matrix for classifier using early fusion with MI-selected features

This study shows how important early feature fusion and MI-based feature selection are for improving all the classifiers' performance, especially for XGBoost. The classifier receives a more condensed and discriminative input by successfully combining several feature representations and removing unnecessary features, eventually increasing the prediction potential for seizure detection.

5. Conclusion and Future Scope

This study examined four feature approaches in order to systematically evaluate an XGBoost classifier for epileptic episode diagnosis using EEG signals. By eliminating redundancy, MI-based selection greatly enhanced the performance of the first HHT features. Early feature fusion (RICA, PCA, EMD, and HHT) led to further developments, emphasizing the importance of combining dissimilar data.

Accuracy was consistently greatly increased by MI-based selection and early feature fusion. Combining these two methods eventually produced the highest accuracy, proving that a seizure detection system may be made extremely accurate and successful by combining a variety of variables and carefully choosing the most relevant ones. While acknowledging its current dataset-specific performance, future research will test this technique on bigger, more diverse EEG datasets to confirm its broad applicability and reduce any biases.

References

[1] Epilepsy, World Helath Organization, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/epilepsy

[2] Mădălina-Giorgiana Murariu, Florica-Ramona Dorobanțu, and Daniela Tărniceriu, "A Novel Automated Empirical Mode Decomposition (EMD) Based Method and Spectral Feature Extraction for Epilepsy EEG Signals Classification," *Electronics*, vol. 12, no. 9, pp. 1-20, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Muzaffer Aslan, and Zeynep Alçin, "Detection of Epileptic Seizures from EEG Signals with Hilbert Huang Transformation," *Cumhuriyet Science Journal*, vol. 42, no. 2, pp. 508-514, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Varun Bajaj, and Ram Bilas Pachori, "Separation of Rhythms of EEG Signals Based on Hilbert-Huang Transformation with Application to Seizure Detection," *6th International Conference, Convergence and Hybrid Information Technology*, Daejeon, Korea, pp. 493-500, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[5] Sunkara Mounika, and S.R. Reeja, "Forecasting Epileptic Seizures Using XGBoost Methodology and EEG Signals," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 10, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Hepseeba Kode, Khaled Elleithy, and Laiali Almazaydeh, "Epileptic Seizure Detection in EEG Signals Using Machine Learning and Deep Learning Techniques," *IEEE Access*, vol. 12, pp. 80657-80668, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Kazi Mahmudul Hassan et al., "Epileptic Seizure Detection in EEG Using Mutual Information-Based Best Individual Feature Selection," *Expert Systems with Applications*, vol. 193, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Khaled M. Alalayah et al., "Effective Early Detection of Epileptic Seizures through EEG Signals Using Classification Algorithms Based on t-Distributed Stochastic Neighbor Embedding and K-Means," *Diagnostics*, vol. 13, no. 11, pp. 1-24, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Dinesh Kumar Atal, and Mukhtiar Singh, "A Hybrid Feature Extraction and Machine Learning Approaches for Epileptic Seizure Detection," *Multidimensional Systems and Signal Processing*, vol. 31, pp. 503-525, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[10] U. Rajendra Acharya et al., "Automated EEG-Based Screening of Depression Using Deep Convolutional Neural Network," *Computer Methods and Programs in Biomedicine*, vol. 161, pp. 103-113, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[11] Navdeep Kaur, "A Review Paper On Copy Move Forgery Detection Techniques," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 7, pp. 157-161, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[12] Varsha Harpale, and Vinayak Bairagi, "An Adaptive Method for Feature Selection and Extraction for Classification of Epileptic EEG Signal in Significant States," *Journal of King Saud University – Computer and Information Sciences*, vol. 33, no. 6, pp. 668-676, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Larbi Boubchir, Boubaker Daachi, and Vinod Pangracious, "A Review of Feature Extraction for EEG Epileptic Seizure Detection and Classification," *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*, Barcelona, Spain, pp. 456-460, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[14] Abeg Kumar Jaiswal, and Haider Banka, "Local Pattern Transformation based Feature Extraction Techniques for Classification of Epileptic EEG Signals," *Biomedical Signal Processing and Control*, vol. 34, pp. 81-92, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Qi Sun et al., "Automatic Epileptic Seizure Detection Using PSO-Based Feature Selection and Multilevel Spectral Analysis for EEG Signals," *Journal of Sensors*, vol. 2022, no. 1, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Anis Malekzadeh et al., "A Method for Epileptic Seizure Detection in EEG Signals Based on Tunable Q-Factor Wavelet Transform Method Using Grasshopper Optimization Algorithm with Support Vector Machine Classifier," *Internal Medicine Today*, vol. 28, no. 1, pp. 98-127, 2022. [Google Scholar] [Publisher Link]

[17] Guanqing Kong et al., "A Novel Method for Optimizing Epilepsy Detection Features through Multi-Domain Feature Fusion and Selection," *Frontiers in Computational Neuroscience*, vol. 18, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] Abdulhamit Subasi, and M. Ismail Gursoy, "EEG Signal Classification Using PCA, ICA, LDA and Support Vector Machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659-8666, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[19] Ling Guo, Daniel Rivero, and Alejandro Pazos "Epileptic Seizure Detection using Multiwavelet Transform based Approximate Entropy and Artificial Neural Networks," *Journal of Neuroscience Methods*, vol. 193, no. 1, pp. 156-163, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Lina Wang et al., "Automatic Epileptic Seizure Detection in EEG Signals using Multi-Domain Feature Extraction and Nonlinear Analysis," *Entropy*, vol. 19, no. 6, pp. 1-17, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[21] Prasad Sarappadi Narasimha, and B.S. Kavya, “Optimized Feature Extraction Process to Identify the Seizure using Feature Fusion Technique,” *IAENG International Association of Engineers Engineering*, vol. 29, no. 2, pp. 1-16, 2021. [Google Scholar]

[22] Athar A. Ein Shoka et al., “Automated Seizure Diagnosis System Based on Feature Extraction and Channel Selection Using EEG Signals,” *Brain Informatics*, vol. 8, pp. 1-16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] Ly V. Tran et al., “Application of Machine Learning in Epileptic Seizure Detection,” *Diagnostics*, vol. 12, no. 11, pp. 1-26, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Shamriz Nahzat, and Mete Yağanoğlu, “Classification of Epileptic Seizure Dataset Using Different Machine Learning Algorithms and PCA Feature Reduction Technique,” *Journal of Investigations on Engineering & Technology*, vol. 4, no. 2, pp. 47-60, 2021. [Google Scholar] [Publisher Link]

[25] Syed Muhammad Usman, Muhammad Usman, and Simon Fong, “Epileptic Seizures Prediction Using Machine Learning Methods,” *Computational and Mathematical Methods in Medicine*, vol. 2017, no. 1, pp. 1-10, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[26] Kavita Mahajan, M.R. Vargantwar, and Sangita M. Rajput, “Classification of EEG Using PCA, ICA and Neural Network,” *International Conference in Computational Intelligence*, no. 6, pp. 1-4, 2012. [Google Scholar] [Publisher Link]

[27] Quoc V. Le et al., “ICA with Reconstruction Cost for Efficient Overcomplete Feature Learning,” *Advances in Neural Information Processing Systems*, 2011.[Google Scholar] [Publisher Link]

[28] Zhen Jiang, and Wenshan Zhao, “Optimal Selection of Customized Features for Implementing Seizure Detection in Wearable Electroencephalography Sensor,” *IEEE Sensors Journal*, vol. 20, pp. 12941-12949, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[29] Tianqi Chen, and Carlos Guestrin, “XGBoost: A Scalable Tree Boosting System,” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California USA, pp. 785-794, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[30] Marina Sokolova, and Guy Lapalme, “A Systematic Analysis of Performance Measures for Classification Tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[31] Wenna Chen et al., “An Automated Detection of Epileptic Seizures EEG Using CNN Classifier based on Feature Fusion with High Accuracy,” *BMC Medical Informatics and Decision Making*, vol. 23, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]