Dokare and Gupta
Health Information Science and Systems (2025) 13:30
https://doi.org/10.1007/s13755-025-00348-4

Health Information Science and Systems

RESEARCH

Optimized seizure detection leveraging band-specific insights from limited EEG channels

Indu Dokare[1,2*†] and Sudha Gupta[1†]

Abstract

***Purpose*:** Effective seizure detection systems are crucial for health information systems and managing epilepsy, yet traditional multichannel EEG devices can be costly and complex. This study aims to optimize EEG channel selection and focus on specific frequency bands associated with epileptic activity, enhancing the system's usability and accuracy for clinical applications.

***Methods*:** This work proposes a novel method by integrating channel selection with band-wise analysis for seizure detection. The channel selection uses an ensemble of mutual information (MI) and Random Forest (RF) techniques to select the most relevant channels. The signals from the selected channels are decomposed into different frequency bands using discrete wavelet transform (DWT). To evaluate the effectiveness of this approach, ten features are extracted from each frequency band and then classified using a support vector machine (SVM) classifier.

***Results*:** This work has obtained a mean accuracy of 97.70%, a mean sensitivity of 86.70%, and a mean specificity of 99.66% for seizure patients from a well-established CHB-MIT dataset and an almost 80% reduction in processing time.

***Conclusion*:** These benefits make seizure detection devices more wearable, less intrusive, and easier to integrate with other health monitoring systems, allowing for discreet and comfortable monitoring that supports an active lifestyle for patients.

Keywords: Epilepsy, Multichannel EEG signal, Ensemble-based channel selection, Band-wise analysis.

Introduction

Millions of individuals worldwide suffer from epilepsy, a neurological condition marked by spontaneous and recurring seizures [20, 30]. Electroencephalography (EEG), which records brain electrical activity using electrodes placed on the scalp [4], is a useful diagnostic and monitoring tool [1] widely used for epilepsy. Several channels are typically placed on the scalp during EEG recordings to record brain activity from different spatial regions. These multichannel EEG systems provide rich and accurate data, in contrast, these also bring difficulties related to computing complexity, data storage, and poor signal-to-noise ratio. Processing and analyzing such multichannel EEG data becomes computationally intensive in the case of real-time applications [5] or resource-constrained environments [7] such as wearable devices. Noisy channels negatively impact the performance of a system [5] and increase response time due to the complexity introduced by multiple channels.

Intending to reduce the dimensionality of EEG data without sacrificing its diagnostic value, channel reduction approaches have become essential in this context. Manual channel selection in seizure detection offers challenges [27] including subjectivity, expertise dependency, time consumption, inability to adapt, limited scalability, and the risk of overlooking relevant channels. Individual judgment and expertise can cause inconsistency and

†Indu Dokare and Sudha Gupta have contributed equally to this work.

*Correspondence: indu.dokare@somaiya.edu
[1] Department of Electronics Engineering, K. J. Somaiya College of Engineering, Somaiya Vidyavihar University, Vidyanagar, Vidyavihar East, Mumbai, Maharashtra 400077, India
Full list of author information is available at the end of the article

inefficiency, especially in large-scale or real-time monitoring circumstances. On the other hand, an automated channel selection in seizure detection provides objective, efficient, and scalable techniques for detecting relevant brain activity. These algorithms reduce subjectivity and variability, streamlining the detection process and allowing for rapid analysis of big datasets. Automated channel selection uses machine learning and deep learning approaches [14, 29] to continuously enhance accuracy and flexibility over time.

Different frequency bands such as $\gamma, \beta, \alpha, \theta$, and δ are associated with various neurological activities [15, 22, 25], and changes in these bands can provide important information about the occurrence and characteristics of seizures. Certain frequency ranges are more localized to specific areas of the brain. This proposed work aims to include a selection of the channels most significant to seizure occurrence detection and then perform a band-wise analysis to provide a more focused and effective evaluation.

The next part of this paper is structured as follows: Sect. 2 covers existing works, sect. 3 includes materials and methods containing all steps involved in the proposed work such as channel selection methods, filter and segmentation, signal decomposition, feature extraction, normalization, classification, and performance metrics used. The experimental findings along with the discussion are reported in sect. 4 whereas the conclusion is explained in the sect. 5.

Existing works

Various methods are investigated for dimension reduction such as principal component analysis (PCA) [8, 32] which transforms the original multichannel EEG signals into a new set of reduced channels. However, such methods will completely lose the origin of signal generation and can't be effectively used in further analysis in applications such as seizure detection. Hence, channel selection methods provide greater interpretability because the channels chosen directly represent the features contributing to the study.

In the last several years, researchers have investigated several types of channel selection strategies [3] to improve the effectiveness of seizure detection systems. These methods range from traditional approaches based on domain expertise or manual selection [9] to intricate data-driven approaches employing machine learning and signal processing algorithms using scalp EEG [7, 20, 28] or intracranial EEG (iEEG) [12, 27]. In the work [27] based on automatic channel selection, 3 iEEG channels were selected using an ensemble algorithm such as gradient boosting, AdaBoost, and random forest. The sensitivity and specificity obtained were 85.7% and 98.01% respectively. The channels were selected based on the highest variance during seizure [12], where the wavelet analysis was used to extract features and SVM was used as a classifier. This study achieved a sensitivity of 96% using three iEEG channels for seizure detection tasks. Further authors have used t-distributed stochastic neighbor embedding (t-SNE) for dimension reduction. Another study [9] has manually selected channels and demonstrated a decrease in computational load for seizure prediction.

Several studies have employed different channel selection methods using scalp EEG CHB-MIT dataset and the related work is briefed as follows: The variance-based method employed for channel selection [13] has extracted 11 features from selected 3 channels and after averaging, these are fed to the seven classifiers. Using the KNN algorithm, this study achieved an accuracy of 89.02%, a sensitivity of 100%, and a specificity of 77.5% for one patient. Another study [7] has attained the power spectral density of all EEG channels and then the best channels were selected using a random forest algorithm. Next, the authors utilized t-distributed Stochastic Neighbor Embedding (t-SNE) to further reduce the feature dimensionality. KNN was employed as a classifier to classify seizure and non-seizure events. The authors have achieved a sensitivity of 80.87% and a precision of 47.45%. The work proposed [8] examined the performance of the artificial neural network (ANN) on EEG signals by applying PCA for the selection of channels. Out of the 23 channels considered, after using PCA, the highest accuracy of 86.7% is achieved with 18 channels. A non-dominated sorting genetic algorithm (NSGA) was proposed [20] for channel selection in which two energy values and two fractal dimension values from both the EMD and DWT sub-bands were estimated as features. These features were fed to KNN, NB, SVM, and RF classifiers. This work has achieved a percentage accuracy in the range of 77.6 to 100. The proposed work [23] has selected the channel using neighborhood component analysis (NCA), extracted the best features from statistical features, and further classified features using ReliefF-based optimization (RBO) in combination with a k-nearest neighbor classifier and has obtained a sensitivity of 100%. All these reported works are different in terms of the dataset used, type of EEG, and seizure task as detection or prediction.

To summarize, despite the significant advancements in EEG-based seizure detection using limited channels over the years, earlier approaches still face several challenges that limit their effectiveness as illustrated in Table 1. Many of these methods relied on a full set of channels [3] without proper selection, leading to unnecessary computational overhead and potential

Table 1 Summarization of existing works

Works	Channel reduction method	EEG type/NC	Dataset/NP	Limitations
[9]	Manual	sEEG/6	CHB-MIT/6	Time-consuming and prone to human error. A poor generalization.
[12]	Variance	ECoG/3	FHS/10	May fail to handle non-linear relationships between channels, and no relatonship between channels and the target variable.
[32]	PCA	sEEG/3-20	CHB-MIT/6	Poor accuracy, unsupervised method, may fail to account for the relationship between channel data and the target variable.
[7]	RF	sEEG/3	CHB-MIT/23	Poor sensitivity, and poor generalization.
[27]	Ensemble classifiers	iEEG/3	Kaggle/8	A single method may lack generalization.
[23]	NCA	sEEG/8	CHB-MIT/6	A single method may face challenges in generalization.
[20]	NSGA	sEEG/2	CHB-MIT/24	Lack of generalization.
[13]	Variance	sEEG/3	CHB-MIT/8	May fail to handle non-linear relationships, and no relationship between channels and the target variable.
[2]	NN with attention mechanism	sEEG/3	CHB-MIT/24	Poor accuracy and sensitivity.
Proposed work	**Ensemble of MI and RF**	**sEEG/5**	**CHB-MIT/24**	**Good generalization. Considers relationship between channels and target variable providing robust performance.**

The proposed work is highlighted in bold

NC number of channels, *NP* number of patients

redundancy. Additionally few traditional systems [9] typically depended on manual channel selection, which was both time-consuming and prone to human error. While some studies [8, 32] employed PCA for dimensionality reduction, this approach had its limitations. PCA, being an unsupervised method, fails to account for the relationship between channel data and the target variable, potentially discarding important seizure-related information. Even PCA transforms the original sets of channels into a new set of uncorrelated components, which may not always correspond to physiologically meaningful patterns, making the results harder to interpret. Even the variance-based methods [13] can't handle the non-linear relationship between channels.

Furthermore, most of the works [2, 7, 8] have employed a single method for channel selection. Using a single method for channel selection can lead to limitations in capturing diverse aspects of data, potentially introducing bias and reducing the robustness of the seizure detection system. This approach may not fully exploit the strengths of different algorithms, limiting the system's ability to generalize across different datasets and patient populations.

To address these challenges and seizure pattern differences in patients, our proposed work emphasizes the importance of channel selection in the context of seizure detection and has incorporated the following approaches:

- This proposed work has employed an ensemble of mutual information (MI) ranking and random forest (RF) based channel ranking for channel selection. This ensemble approach makes a decision based on the channel ranking results obtained from two models, due to this it provides a more accurate set of channels to be selected. Both methods can complement each other, as one method may miss certain features that the other captures, enabling accurate detection of pattern variability across patients and can provide results with greater generalization.
- Seizure pattern variability is observed within a patient and even among patients. To capture such variations highly adaptive methods are required. MI can detect complex, non-linear dependencies between EEG channels and doesn't assume any specific relationship form, making it highly adaptive to patient-specific seizure characteristics. MI is sensitive to capturing subtle differences, which allows it to adapt to the variability of seizure patterns observed in different patients. Since RF uses multiple decision trees which makes it robust in identifying and learning from these variable seizure patterns. Hence, the ensemble of MI and RF methods complement each other, effectively capturing and addressing patient differences.
- Additionally this work proposes the approach for a targeted assessment of EEG data to identify relevant patterns associated with epileptic occurrences in different frequency bands.
- Furthermore, the employment of the statistical and entropy-based feature extraction of selected channels aids in capturing the diverse seizure patterns.

Relying on a single method may not fully capture the complexity of the data or meet clinical demands. A dual-opinion approach, integrating decisions from methods like MI and RF, ensures that the final set of channels is both statistically robust and clinically relevant. Such a method not only improves diagnostic accuracy but also facilitates faster computation, greater interpretability, and better generalization, making it an essential step toward practical and reliable medical applications. This proposed method improves decision-making in clinical practice and research, making it an important tool for channel selection in wearable devices.

Materials and methods

This proposed work aims to construct an accurate seizure detection system with less computational load, better response time, and good sensitivity. The major contribution of this work is to provide an effective channel selection method along with a band-wise analysis of EEG signals for seizure detection. Figure 1 shows the framework of this proposed approach and the steps undertaken are summarized as follows:

1. *Channel selection*: The channel selection module comprises an ensemble of channel ranking based on mutual information and random forest to select 'k' channels from 22 EEG channel recordings of 24 patients using the CHB-MIT dataset.
2. *Filtering and segmentation*: The data from the selected channels is applied to the Butterworth low pass filter and notch filter to remove the noise signals. This filtered signal is then segmented in a fixed length segment of 4 s.
3. *Signal decomposition*: To obtain the frequency bands of interest, the signal is decomposed using DWT obtaining five wavelet coefficients corresponding to the five frequency bands such as δ, θ, β, α, and γ bands.
4. *Feature extraction*: The statistical and entropy-based features are estimated from each wavelet coefficient.
5. *Data normalization*: Z-score normalization is employed to standardize the features by making the mean zero and scaling them to unit variance.
6. *Classification*: A set of statistical and entropy-based features of every wavelet coefficient is applied to a widely used classifier, support vector machine (SVM).
7. *Cross-validation*: To assess the performance of SVM classifier model, a 5 fold cross-validation technique is employed.
8. *Performance metrics*: The performance of the classifier model is evaluated using accuracy, sensitivity, and specificity.

Automatic channel selection

When employing channel selection for multichannel scalp EEG data for seizure detection, less important channels are excluded, retaining just the most informative subset of EEG channels. Finding a subset of channels that offer pertinent data for identifying and characterizing seizures while reducing noise and irrelevant signals is the key to effective EEG channel selection for seizure detection. The primary goal of channel reduction method is to minimize computational complexity and cost while addressing seizure variability, ensuring generalization, and retaining essential features. This model determines a measure of each channel's importance and how much each channel contributes to the accuracy of the model. This proposed work has used an ensemble of channel rankings obtained by mutual information (MI) based and random forest (RF) based methods for channel selection. It employs a novel approach to select the final set of channels by integrating two subsets derived from MI and RF-based methods. This strategy aids in identifying the most informative EEG channels while simultaneously minimizing computational complexity, making a more robust and generic method. The channels are ranked according to the importance score. A higher importance score is assigned to the channel, leading to more accurate predictions. After ranking the channels according to their metrics using both methods, 'k' channels are selected using the ensemble method for further analysis. This proposed work has considered 22 EEG channels such as 'FP1-F7', 'F7-T7', 'T7-P7', 'P7-O1', 'FP1-F3', 'F3-C3', 'C3-P3', 'P3-O1', 'FP2-F4', 'F4-C4', 'C4-P4', 'P4-O2', 'FP2-F8', 'F8-T8', 'T8-P8-0', 'P8-O2', 'FZ-CZ', 'CZ-PZ', 'P7-T7', 'T7-FT9', 'FT9-FT10', and 'FT10-T8' from the CHB-MIT dataset.

Mutual information-based channel selection

The data-driven technique based on MI for channel selection allows the computer to infer information from the EEG data itself, rather than making explicit assumptions about the importance of individual channels. Mutual information (MI) helps to identify the most informative and non-redundant EEG channels by measuring how strongly each channel relates to seizure patterns, ensuring that only the most relevant ones are selected. MI measures the dependency or shared information between two variables, such as an EEG channel and the target variable (e.g., seizure vs. non-seizure). This allows them to identify channels that are most relevant to the

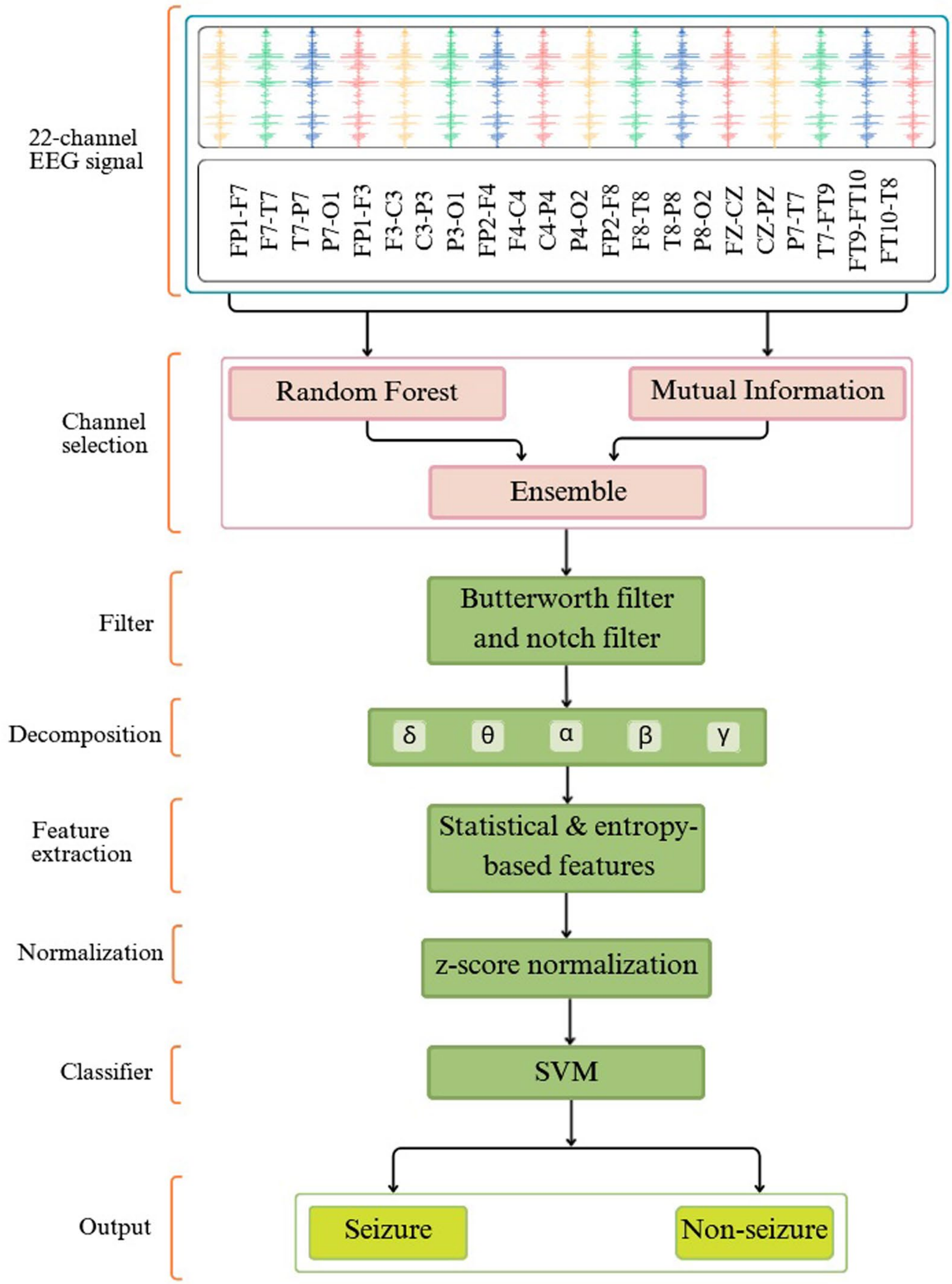

Fig. 1 Proposed framework for seizure detection: this figure outlines the stages of the proposed system, including ensemble method of channels selection, data pre-processing, signal decomposition, feature extraction, data normalization, and the classification stage

specific characteristics of the data, effectively addressing differences in signal patterns across patients. It is an ideal tool for EEG channel selection due to its ability to handle noisy, complex, and high-dimensional data. This method is excellent for explicitly identifying the dependency between channels and target labels, capturing linear and non-linear relationships unlike correlation, which only captures linear relationships. This is particularly important because seizure patterns can vary widely among patients and may not follow simple linear trends. Seizure patterns in different patients often exhibit non-linear dynamics, which MI effectively identifies. This ensures that even subtle or complex dependencies between EEG channels and seizure events are not overlooked. MI measures how much knowing the value of one variable reduces uncertainty about the other by quantifying the statistical dependency between two random variables, in this case, the EEG channels and the target variable. MI of two random variables X and Y is a measure of the mutual dependence between the two variables. It measures the statistical dependence between each channel and the labels as target variables as given by Eq. 1 [11].

$$I(X;Y) = \sum_{x \epsilon X} \sum_{y \epsilon Y} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) \quad (1)$$

where $p(x,y)$ is the joint probability mass function of X and Y, $p(x)$ and $p(y)$ are the marginal probability mass functions of X and Y respectively.

After computing the MI between each channel and the target variable, the channels are ranked according to the estimated MI values. The target variable is labeled as '1' for seizure and '0' for non-seizure state. A high MI value indicates a stronger association between the channel and the target variable. The MI values obtained for the channels of a patient vary, determining the channel rankings for that patient. Furthermore, the channel rankings differ across patients, indicating that MI effectively identifies patient-specific variations.

Tree-based channel selection

The RF, a tree-based method is used to identify the most informative EEG channel for a certain task such as seizure detection. This implementation has utilized the sci-kit-learn library to implement RF for EEG channel ranking. RF is an ensemble learning method that uses multiple decision trees [6] during training to build a model and combines their predictions to increase accuracy and robustness [18]. Each decision tree is built from a subset of the training data considering channels as input features in this case. Furthermore, randomization is used throughout tree development, such as random channel selection at node splits. This randomization helps to decorrelate the separate trees, reducing overfitting [18]. The channel importance is determined based on how much each EEG channel contributes to decreasing impurity in the trees. This channel selection method has employed the Gini impurity to assess the channel importance score. Essentially, it measures the decrease in impurity caused by each channel when it's used in a node split across all the trees in the forest. Features that lead to large impurity decreases are considered more important.

The tree-based structures, like those used in RF, inherently help in identifying patient-specific differences due to their hierarchical decision-making process. Each decision tree in the RF model splits data using the most informative channels, adapting to the unique patterns and distributions in each patient's EEG signals and seizure patterns. Since seizure patterns and EEG characteristics differ from one patient to another, the tree-based structure naturally identifies these differences by adapting the splits and importance scores to patient-specific variations. This flexibility enables tree-based methods to capture complex, non-linear relationships within the data, making them highly effective for detecting seizures in different patients. By adapting to these differences, RF effectively tailors the channel rankings and selection to each patient's data, helping to capture patient-specific variations in seizure detection.

EEG data is properly formatted with channels as features and seizure presence or absence as the target variable. This implementation has used the RandomForestClassifier class from sci-kit-learn to train the RF model on a 22-channel EEG signal. After the model is trained, the feature importance attribute is accessed to retrieve the importance scores for each EEG channel. Then the channels are ranked based on it's importance score. Channels with higher importance scores are considered more informative for seizure detection.

Ensemble of channel ranking

This research work has proposed a novel method for selecting the channels based on ensemble of MI and RF based methods. Both, MI and RF-based methods capture complex and non-linear relationships between the EEG channels. MI-based methods adapt by evaluating the relationship between EEG channel signals and seizure patterns (labels) within each dataset. This allows them to identify channels that are most relevant to the specific characteristics of the data, effectively addressing differences in signal patterns across patients. RF-based methods implicitly handle redundancy by focusing on the most impactful channel during splits. MI-based methods are excellent for explicitly identifying the dependency between channels and target labels, capturing linear and non-linear relationships. In contrast, RF-based methods are more practical, offering adaptability across datasets, easier interpretability, and the ability to implicitly manage redundant channels during model training.

Hence, in this proposed work of channel selection, we have combined both methods called ensembling which leverages the strengths of both methods for more accurate and efficient channel selection. The top common channels are selected by using intersection operation. The intersection operation identifies common elements across datasets, ensuring consistency, reducing noise, and streamlining analysis. It enhances feature selection, decision-making, and resource efficiency, making it essential for reliable and focused data processing. The main objective of ensembling MI and RF methods are to utilize their complementary strengths for better generalization and robustness, especially in diverse and noisy datasets. Additionally, it increases confidence in the selected channels, as channels identified by both methods are likely to be highly relevant. This comprehensive approach reduces bias and enhances the generalization capability of the seizure detection model, making it more effective across various datasets and patient populations. Overall, combining multiple methods leads to a more reliable, accurate, and efficient channel selection method for a seizure detection system used for a diverse population.

As shown in algorithm 1, after ranking channels using both channel selection methods, the top 'k' channels are selected by an intersection operation. The value of 'k' ranges from 1 to 21, since this study has used 22 EEG channels. The value of 'k' can be selected according to the need of the study.

Algorithm 1 Ensemble of two channel ranking: top k channel selection

```
procedure TOPK_COMMON_CHANNELS(list1, list2)
    common_channels ← ∅
    combined_ranks ← ∅

    for ch ∈ list1 do
        if ch ∈ list2 then
            Add ch to common_channels
        end if
    end for

    for ch ∈ common_channels do
        rank1 ← index of ch in list1
        rank2 ← index of ch in list2
        combined_rank ← rank1 + rank2
        combined_ranks[ch] ← combined_rank
    end for

    sorted_common_channels ← sort keys of combined_ranks by values in ascending order
    top_k_common_channels ← first k elements of sorted_common_channels

     return top_k_common_channels
end procedure
```

Filtering and segmentation

Filtering is essential to focus on the specific frequency range of interest required for epilepsy analysis. The EEG signal is filtered with a 4th order low-pass Butterworth filter with a lower cutoff frequency of 64 Hz and further, it is notch filtered with a 60 Hz frequency to eliminate a 60 Hz line interference. A low-pass filter will remove high-frequency noise components present in the signal. The filtered signal is segmented in a fixed length of 4 s, which is not too small like 1 s [19], or high such as 6 or 30 s [31].

Signal decomposition using DWT

The seizure activity can be prominently seen in a specific frequency band for a particular patient. It can be observed in lower-frequency bands or higher-frequency bands. This prominence of seizure activity in a specific frequency band [22] would help to understand the underlying seizure activity very well and contribute to the diagnosis and management of epilepsy. This study leverages DWT to decompose EEG signals into different frequency bands providing multi-resolution analysis. DWT decomposes the signal by passing the signal through a series of high-pass and low-pass filters, obtaining the approximation and detail coefficients at different levels. In this work, Daubechies wavelet family, db4 is used and the signal is decomposed to level 4. The obtained wavelet coefficients cA4, cD4, cD3, cD2, and cD1 are used for further processing. The approximation coefficients cA4 correspond to the δ band (0–4 Hz), and detail coefficients cD4, cD3, cD2 and cD1 represent θ band (4–8 Hz), α

band (8–16 Hz), β band (16–32 Hz) and γ band (32–64 Hz) respectively.

Feature extraction

Different sets of features help to address patients' seizure pattern variability by capturing diverse aspects of EEG signal characteristics unique to each patient. EEG signals can exhibit significant differences in frequency, amplitude, and patterns of brain activity. In this proposed work, enhanced interpretability of EEG signals is achieved through feature fusion, where raw signals are transformed into more insightful representations [17] by estimating statistical and entropy-based features as defined in Table 2. These features are extracted from every coefficient cA4, cD4, cD3, cD2 and cD1. The statistical features such as the minimum, maximum, mean, standard deviation, variance, skewness, and kurtosis of the coefficients are determined according to equations 2 through 8. Along with these seven features, three entropy-based features such as sample entropy, permutation entropy, and Shannon entropy of the coefficients are estimated using equations 9 through 11. Hence, a total of ten features are extracted from each coefficient. By combining and analyzing these distinct features, the variability in seizure patterns and signal characteristics across patients can be better represented. This ensures the model adapts to a diverse patient-specific traits, improving its robustness and generalizability. Further, these ten extracted features are fed to the SVM classifier.

Classification

SVM is a powerful supervised learning algorithm that finds the optimal hyperplane maximizing the margin between classes, using support vectors to define this boundary [10]. After labeling the seizure segment features as '1' and non-seizure segment features as '0', a feature vector of ten features is applied to a support vector machine (SVM) classifier. The data is split into the training set constituting 70% of the total dataset while the remaining 30% of the dataset is allocated for the testing set. This proposed work has employed a grid search method with 5-fold cross-validation to set the parameters of SVM. Grid search is a systematic method for tuning SVM parameters to find the optimal combination that yields the best results for a given dataset. Key parameters for SVM tuning include regularization parameter 'C' which balances model complexity and training accuracy and 'gamma' which controls the influence of individual data points for the non-linear kernel. Since EEG signals are highly complex and non-linear, we opted for a non-linear SVM with an RBF kernel. We have set 'C' values as 1, 10, and 100 along with 'gamma' values as 0.001 and 0.0001 to fine-tune the SVM model. These values are selected to balance model flexibility and generalization. 'C' values controls the trade-off between margin size and misclassification tolerance, with smaller 'C' allowing

Table 2 Description of features used in this proposed work

Sr.no.	Features	Description	Equation no.
1	Minimum	$Min = X_{min} = min\{x_1, x_2, \ldots, x_n\}$	(2)
2	Maximum	$Max = X_{max} = max\{x_1, x_2, \ldots, x_n\}$	(3)
3	Mean	$Mean = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	(4)
4	Variance	$Variance = \frac{\sum(x_i-\bar{x})^2}{n}$	(5)
5	Standard deviation	$S = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n}}$	(6)
6	Skewness	$g_1 = \frac{m_3}{m_2^{\frac{3}{2}}}$ where $m_3 = \frac{1}{n}\sum_{i=1}^{n}\left((x_i-\bar{x})^3\right)$ and $m_2 = \frac{1}{n}\sum_{i=1}^{n}\left((x_i-\bar{x})^2\right)$	(7)
7	Kurtosis	$Kurtosis = \frac{1}{n}\frac{\sum_{i=1}^{n}(x_i-\bar{x})^4}{S^4}$	(8)
8	Sample entropy	For a time-series data set of length $N = \{x_1, x_2, x_3, \ldots, x_N\}$ a template vector of length m, such that $X_m(i) = \{x_i, x_{i+1}, \ldots, x_{i+m-1}\}$ and the distance function $d[X_m(i), X_j](i \neq j)$ is to be the Chebyshev distance then, $E_s = -\ln\frac{a}{b}$ a and b are number of template vector pairs having $d[X_{m+1}(i), X_{m+1}(j)] < r$ and $d[X_m(i), X_m(j)] < r$ respectively, where m is embedding dimension, r is tolerance and N are number of data points.	(9)
9	Permutation entropy	For each time series, let p be a probability distribution associated with it where π_i are the frequencies associated with the i possible permutation patterns, therefore $i = 1, 2, \ldots, D!$ where D is the embedding dimension $E_p = -\sum_{i=1}^{D!}\Pi_i \log_2 \Pi_i$	(10)
10	Shannon entropy	$E_{sh}(X) = -\sum_{i=1}^{N} p(x_i)\log_2 p(x_i)$ where $p(x_i)$ is the probability occurrence of feature values from x_1 to x_N	(11)

a wider margin and larger 'C' enforcing stricter boundaries. 'Gamma' defines the influence range of individual data points, with smaller 'gamma' leading to smoother decision boundaries and larger 'gamma' capturing localized patterns. This range ensures a thorough exploration of the model's performance across different regularization and kernel parameters.

The 'C' and 'gamma' parameters are systematically evaluated for all combinations using the grid search method by splitting the dataset into five subsets, where four subsets are used for training and one for validation in each iteration. By ensuring every data point is used for testing exactly once, this method reduces the risk of overfitting and provides a reliable estimate of model performance. The parameter combination that achieves the highest accuracy across all folds has been selected as the optimal configuration in each case, resulting in a robust and well-tuned SVM model.

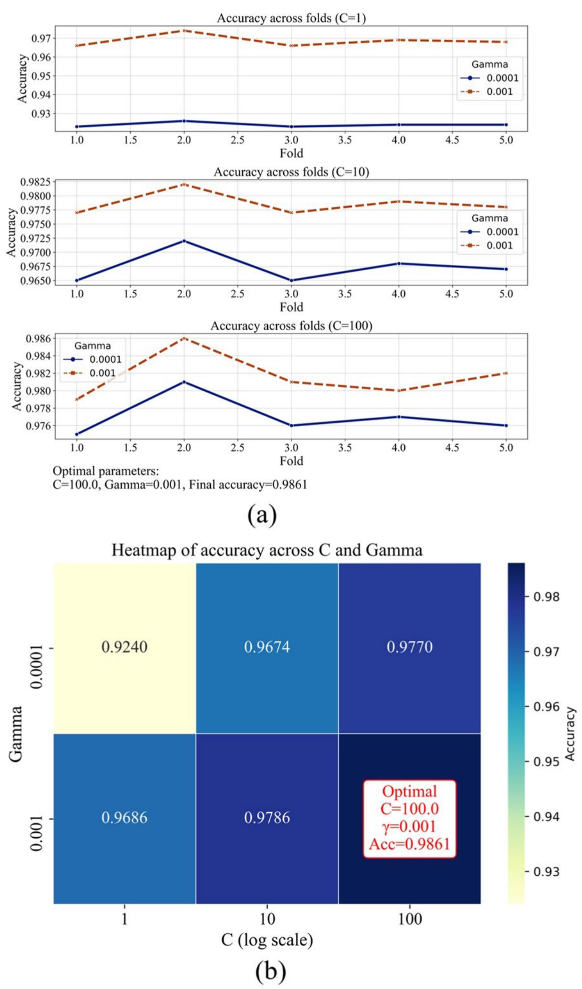

Fig. 2 SVM parameter tuning and and identification of optimal parameters for maximum performance. **a** Accuracy across folds for different values of C and gamma, illustrating the model's performance consistency during cross-validation, **b** heatmap showing the mean accuracy across combinations of 'C' and 'gamma', highlighting the optimal parameters (C = 100, gamma = 0.001) with the highest accuracy of 0.9861

Figure 2 provides a systematic analysis of model performance for varying combinations of the regularization parameter 'C' as 1, 10, and 100 while the kernel parameter 'gamma' as 0.001, and 0.0001. The line plots shown in Fig. 2a illustrate accuracy trends across multiple folds, showing that higher values of 'C' result in improved accuracy and reduced variance, indicating better generalization. Similarly, the choice of 'gamma' significantly impacts performance, with gamma = 0.001 demonstrating superior results compared to gamma = 0.0001 across all 'C' values. The heatmap as shown in Fig. 2b consolidates these findings, identifying C = 100 and gamma = 0.001 as the optimal parameters, achieving the highest accuracy of 0.9861. This systematic process of parameter tuning ensures that the model's performance is optimized by selecting the most effective combination of 'C' and 'gamma'. The optimal values of 'C' and 'gamma' vary from patient to patient due to individual differences in EEG signal patterns.

Performance metrics

The effectiveness of the classifier is evaluated using three performance metrics: accuracy, sensitivity, and specificity, as presented in Table 3. The confusion matrix is used to calculate accuracy, sensitivity, and specificity, which summarizes the predicted versus true values. The overview of components of the confusion matrix is as follows:

- True positive (TP): These are the seizure segments that are correctly identified as seizures by the model.
- True negative (TN): These are the non-seizure segments that are correctly identified as non-seizures by the model.
- False positive (FP): These are the non-seizure segments that are incorrectly identified as seizures by the model.

Table 3 Performance metrics used in this proposed work

Sr. no.	Performance metrics	Equation no.
1	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$	(12)
2	$Sensitivity/Recall/TPR = \frac{TP}{TP+FN}$ where TPR is a true positive rate	(13)
3	$Specificity/selectivity/TNR = \frac{TN}{TN+FP}$ where TNR is a true negative rate	(14)

- False negative (FN): These are the seizure segments that are incorrectly identified as non-seizures by the model.

In seizure detection, accuracy, sensitivity, and specificity are critical performance metrics because they help to evaluate how well a model identifies seizures while minimizing false positives and negatives.

1. Accuracy: It measures the overall correctness of the model by considering both true positives (correct seizure segment detections) and true negatives (correct non-seizure segment detections) as illustrated in equation 12. In seizure detection, a high accuracy indicates that the model is generally making the correct prediction across both seizure and non-seizure periods. However, accuracy alone may not be sufficient, especially when the data is imbalanced.
2. Sensitivity: It is defined as the proportion of actual seizure segments that are correctly identified as seizures by the model as shown in equation 13. It indicates the ability of the model to correctly detect seizures. High sensitivity means that the model is good at detecting seizures, with fewer seizures being missed (false negatives).
3. Specificity: It is defined as the proportion of actual non-seizure segments that are correctly identified as non-seizures by the model as detailed in equation 14. It measures the ability of the model to correctly classify non-seizure periods and avoid false alarms (i.e., incorrectly detecting seizures during non-seizure periods). High specificity means that the model correctly identifies most non-seizure segments and does not raise false alarms during non-seizure periods.

The importance of selecting these three parameters is that sensitivity ensures that the model does not miss actual seizures, addressing safety and intervention needs. Whereas, specificity ensures that the system is not over-sensitive, reducing false alarms and enhancing its usability. Accuracy gives a broader view of the model's performance but needs to be complemented by sensitivity and specificity to understand the trade-offs between detecting seizures and avoiding false alarms.

Results and discussion

This section presents the results of experiments performed on the CHB-MIT dataset of 24 epileptic patients. Each patient dataset is regarded as a distinct experiment, since each patient's seizure pattern is distinct.

Dataset used

The multichannel scalp EEG database (CHB-MIT) [24] used in this work is accessible online at physionet.org [16, 21]. This database was collected at the Children's Hospital Boston (CHB) and the Massachusetts Institute of Technology (MIT) created and contributed this database to Physionet. The EEG signal was recorded by the international 10–20 electrode system and the signals were captured at 256 samples per second with a resolution of 16 bits.

The dataset contains recordings of 5 males and 17 females, ages ranging from 1.5 to 22 years. It consists of 24 cases in total as shown in Table 4. Case chb21 was recorded 1.5 years after case chb01, from the same female patient. Case chb24 was added later in this dataset. The age and gender information of case chb24 is not available in the dataset. A folder per patient contains a continuous recording of 9 to 42 EDF files. The dataset contains 198 seizures in total within 141 distinct files.

Table 4 Dataset used in this proposed work

Sr. no.	Patient ID	Gender	Age (years)	Number of seizures	Seizure duration (s)
1	Patient 1	F	11	7	442
2	Patient 2	M	11	3	172
3	Patient 3	F	14	7	402
4	Patient 4	M	22	4	378
5	Patient 5	F	7	5	558
6	Patient 6	F	1.5	10	153
7	Patient 7	F	14.5	3	325
8	Patient 8	M	3.5	5	787
9	Patient 9	F	10	4	276
10	Patient 10	M	3	7	447
11	Patient 11	F	12	3	806
12	Patient 12	F	2	40	1475
13	Patient 13	F	3	12	828
14	Patient 14	F	9	8	169
15	Patient 15	M	16	20	1992
16	Patient 16	F	7	10	69
17	Patient 17	F	12	3	293
18	Patient 18	F	18	6	317
19	Patient 19	F	19	3	236
20	Patient 20	F	6	8	294
21	Patient 21	F	13	4	199
22	Patient 22	F	9	3	204
23	Patient 23	F	6	7	424
24	Patient 24	–	–	16	511

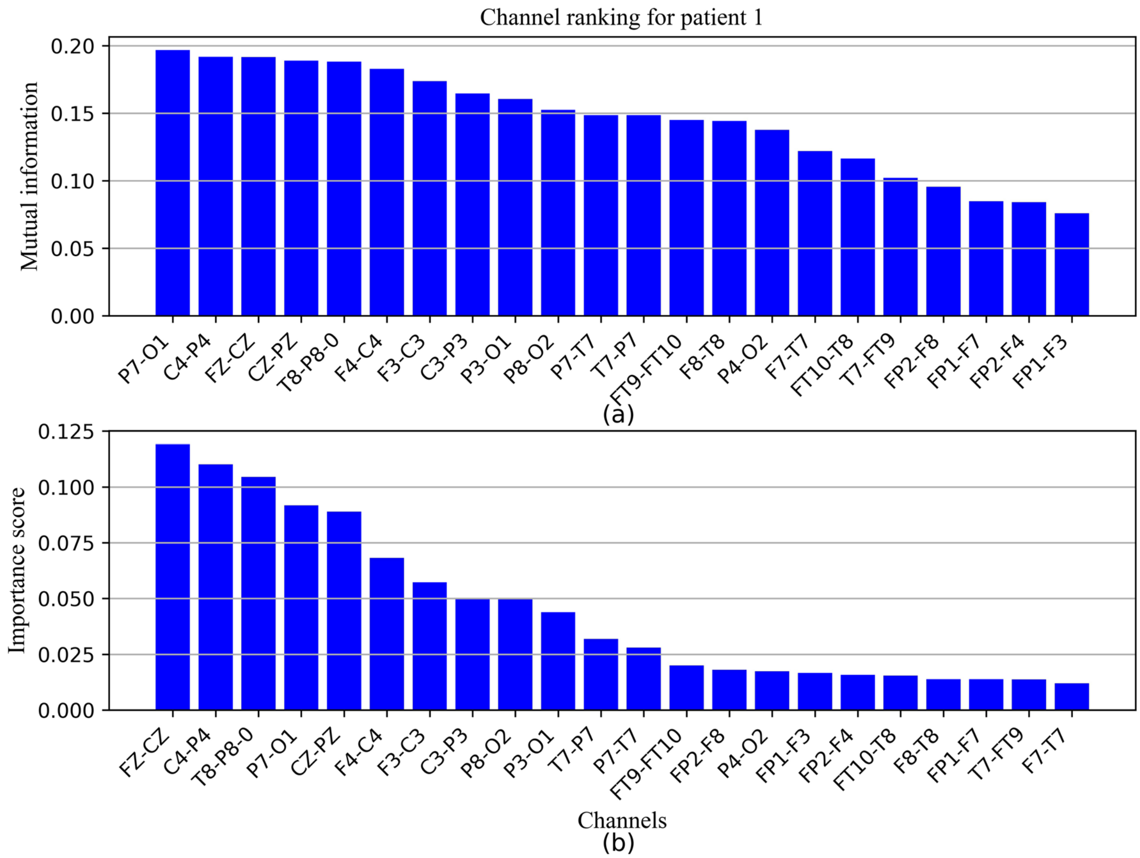

Fig. 3 Channel ranking obtained for patient 1. **a** Channel ranking based on mutual information, **b** channel ranking based on random forest method

Results of channel selection

We utilized the MI model to process data from 22 EEG channels for each patient, obtaining the corresponding MI values, which quantify the dependency between each channel and the target variable. Similarly, we employed the RF model on data from 22 EEG channels for each patient, yielding importance scores to identify the most relevant channels contributing to the target variable. The channel rankings obtained by both methods for patient 1 are shown in Fig. 3. Mutual information obtained for channels indicates a steady decrease, so a change in data applied to this method may change the ranking by one or two channels. Figure 3a depicts a plot of MI value for each channel, with channel 'P7-O1' ranked first and 'FP1-F3' last. The MI values obtained for the channels of a patient vary, determining the channel rankings for that patient. Furthermore, the channel rankings differ across patients, indicating that MI effectively identifies patient-specific variations.

However, the channel importance score estimated using the RF method changes rapidly as shown in Fig. 3b. Channel 'FZ-CZ' has ranked first while channel 'F7-T7' has ranked last. There is relatively little difference in the channel ranking obtained with both approaches. Channel 'P7-O1' is ranked first using the MI-based method, but fourth using the random forest method. The channel 'C4-P4' has ranked second using both methods. Since EEG signals and seizure patterns vary significantly from one patient to another, the importance scores for channels also vary for each patient.

This variation in importance score reflects the unique signal patterns and seizure characteristics of individual patients. By adapting to these differences, RF effectively tailors the channel rankings and selection to each patient's data, helping to capture patient-specific variations in seizure detection. The reason for employing two techniques is to validate one another and capture the variability in the seizure patterns.

After ranking the channels using MI and RF methods, the proposed ensemble approach selects the top 'k' channels. The choice of number of channels (selected 'k' channels) is tailored to the study's goal, ensuring that the system can deliver reliable results while meeting the constraints and objectives of the research. Hence, as proposed in our work, the number of channels ('k') can be selected according to the clinical requirements or patient variation. The trade-off for the number of channels involves balancing processing time, computational complexity, and the performance of seizure detection.

Performance across a different number of channels

We have carried out an investigation in which the impact of altering the number of channels on the model's performance is recorded. The model is trained by altering the number of top channels (top 1 channel, top 2 channels, top 3 channels up to all 22 channels) obtained after ranking by our proposed method across five frequency bands. The performance of a model is estimated using accuracy, sensitivity, and specificity. The confusion matrix as demonstrated in Fig. 4 highlight the model's performance with 5 channels in the δ band for patient 1. It shows 627 non-seizure segments are correctly identified as non-seizures (True Negatives), while 170 seizure segments are accurately detected (True Positives). Only 2 non-seizure segments are mistakenly classified as seizures (False Positives), and just 1 seizure segment is missed (False Negative). The model demonstrates its effectiveness in detecting seizure segments with accuracy of 99.63%, specificity of 99.68%, and sensitivity of 99.42%.

Figure 5 illustrates the performance metrics obtained across five bands for patient 1. It is observed that model's performance is better when top 2, 3, 4, 5, 6 and 7 channels

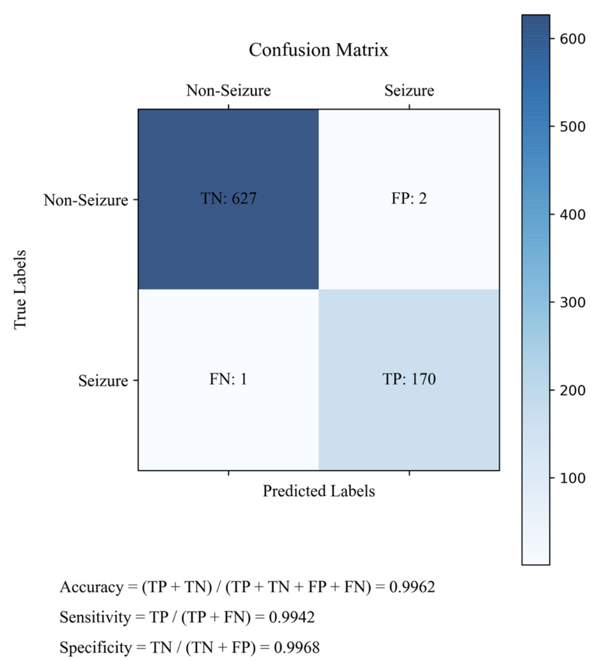

Fig. 4 Confusion matrix obtained for patient 1. It illustrates the classification performance with 5 channels in the δ band, including the estimation of key parameters such as accuracy, sensitivity, and specificity. The matrix highlights the distribution of true positive, true negative, false positive, and false negative classifications

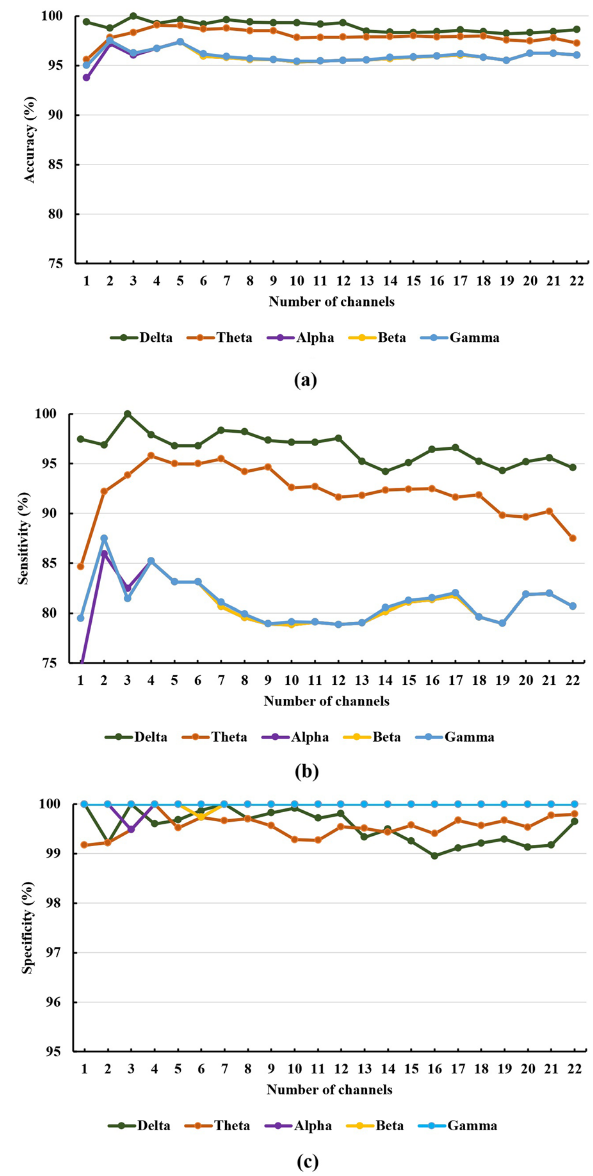

Fig. 5 Band-wise performance metrics obtained for a different numbers of channels for patient 1: **a** band-wise accuracy, **b** band-wise sensitivity, **c** band-wise specificity

are utilized compared to more channels irrespective of frequency bands. The sensitivity obtained for these channels is higher compared to others. The performance of the model for patient 1 in the δ frequency band is superior compared to other bands across all 22 channels.

Processing time across a different number of channels

This proposed work has also assessed the processing time required for the feature extraction and detection stages to investigate the impact of the number of channels on the processing time required. Figure 6a shows

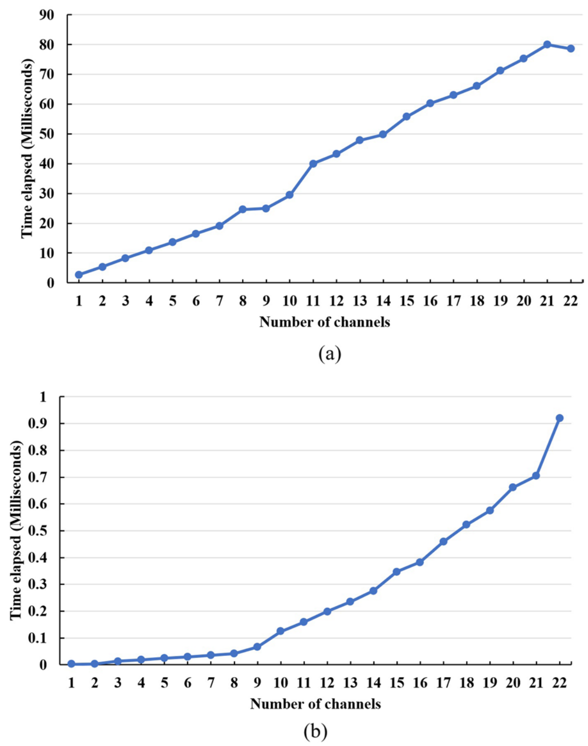

Fig. 6 Processing time elapsed by using a different numbers of channels. **a** Time taken by feature extraction stage, **b** time taken by detection stage

the processing time required for the feature extraction stage as a function of the number of channels selected in the δ band. It shows the linear relationship between the number of channels employed and the processing time required. Figure 6b illustrates the variation in execution time taken by the detection stage as a function of the number of channels selected by our methodology. The detection stage's time grows exponentially with the number of channels employed as shown in Fig. 6b. Hence, as shown in Fig. 6b, the processing time required for the detection stage starts increasing exponentially after the employment of top 8 channels. An extremely low number of channels, such as two, or a high number, such as nine, may not be suitable for achieving optimal performance. These represent boundary conditions where either the information may be insufficient (in the case of too few channels) or diluted and computationally expensive (in the case of many channels) as illustrated in Fig. 6b. Considering this, the performance of a model for the top 3, top 5, and top 7 channels is evaluated. Based on the performance across these three cases, the best-ranked channel is determined for the dataset under analysis.

Performance using the top 3, top 5, and top 7 selected channels

The performance of the model for patient 1 with 22 original channels, the top 3, top 5, and top 7 selected channels is illustrated in Fig. 7. As shown in Fig. 7, the accuracy and specificity obtained are almost identical, while the sensitivity obtained with top 5 channels is maximum compared to top 7 and top 3 channels. The accuracy and sensitivity are improved by 1%, and 5% respectively in the δ band when the top 5 channels are employed compared to the original 22 channels. While a similar specificity is observed in both cases. Figure 7 illustrates that seizure activity is predominantly observed in δ and θ bands for patient 1. In nearly all frequency bands, the model's performance with the top 5 channels is superior to that with other channel selections. With fewer channels, there is less noise and irrelevant information, allowing us to focus on the most useful signals.

To provide a generalized view of the model's effectiveness, we have estimated the average performance across all patients of the CHB-MIT dataset which is shown in Fig. 8. The average performance metrics such as average accuracy, average sensitivity, and average specificity obtained using 22 original channels, using the top 3, top 5, and top 7 channels across five bands are demonstrated in Fig. 8. The accuracy and specificity obtained in these three cases are almost similar. Whereas, the sensitivity obtained by using the top 5 channels outperforms other cases. As shown in Fig. 8, the results demonstrate the consistent and robust performance across patients in the CHB-MIT dataset, highlighting the model's capability to generalize effectively and adapt to the diverse EEG patterns observed in a varied patient population. Tables 5 and 6 present the performance obtained for all patients by using the original 22 channels and the top 5 channels. The average performance across all patients using 22 channels and 5 channels as depicted in Table 6 indicates that the performance using 5 channels is improved over 22 channels. To demonstrate the effectiveness of our model, we have compared the average performance metrics obtained by 22 channels as depicted in Fig. 8a, and 5 channels as shown in Fig. 8c. The accuracy increased by 1%, while specificity remained the same. The sensitivity with 5 channels shows an increase of approximately 5% in the θ, α, β, and γ bands, while the δ band exhibits a notable improvement of around 8% in average sensitivity compared to 22 channels as shown in Fig. 8a, c. Hence, an overall improvement of 5% to 8% in sensitivity is observed with the top 5 selected channels.

Band-wise analysis

Following the implementation of channel selection algorithms, the next contribution of this study is to carry out

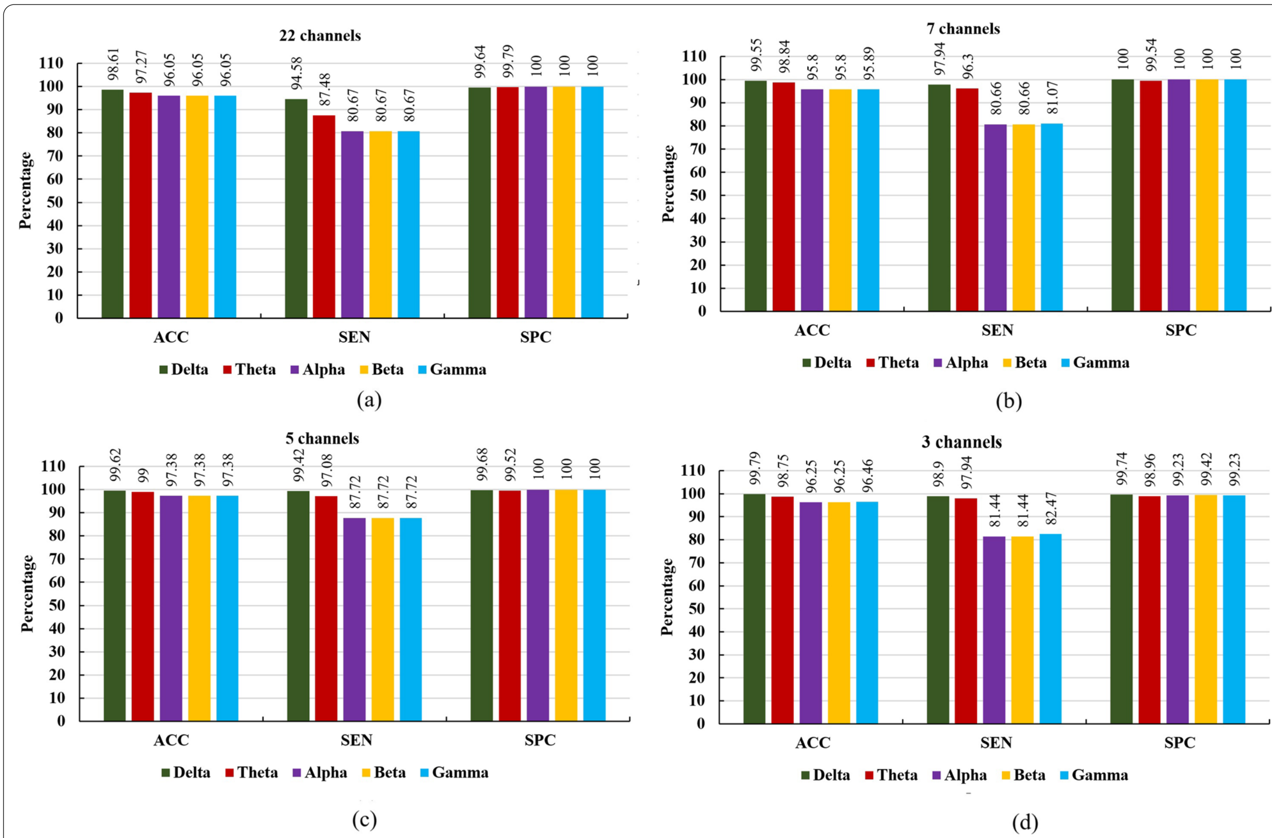

Fig. 7 Performance metrics such as accuracy, sensitivity, and specificity obtained for patient 1. The performance metrics are obtained with different numbers of channels across δ, θ, α, β, and γ bands: **a** original 22 channels, **b** top 7 channels, **c** top 5 channel, **d** top 3 channels

band-wise analysis. Each wavelet coefficient obtained corresponds to a specific frequency range. The coefficients cA4, cD4, cD3, cD2, and cD1 represent the δ, θ, α, β, and γ frequency bands respectively. Each frequency band is individually examined for the seizure detection task in this study. Tables 5 and 6 present the band-wise performance metrics obtained using the original 22 channels and 5 selected channels for each patient.

The system's performance varied among patients, with a few cases such as patients 6, 11, 12, 14, 16, 17, and 21 showing poor sensitivity as reported in Tables 5 and 6. The model achieved good accuracy and specificity (greater than 90%) for these patients. However, sensitivity remained below 70% in these cases. This poor sensitivity could be attributed to short seizure durations in patients 6, 14, and 16. The model may not perform at its best due to a few unusual seizure characteristics or higher noise levels in other cases such as patients 11, 12, 17, and 21. However, it is still capable of detecting seizures with good accuracy and excellent specificity in patients 6, 11, 12, 14, 16, 17, and 21.

Performance across different frequency bands

The performance metrics obtained across five frequency bands are reported in Tables 5 and 6. These metrics are superior in δ and θ bands compared to others in the case of patient 1. The performance of numerous patients, including patients 2, 3, 7, 8, 9, 13, 19, and 20, is almost identical in all frequency bands. Certain frequency bands employing 5 channels show a notable improvement in performance in a subset of patients, including 5, 6, 11, 12, 15, 22, 23, and 24. As shown in Fig. 5, the performance of the model is excellent in the δ band across all channels for patient 1. We found that various frequency bands had varying performance metrics for different patients. This variability can be attributed to individual variations in brain activity and how seizures affect different frequency ranges.

Comparison with other approaches related to channel selection

The main goal of employing channel selection in the seizure detection system is to reduce the computational burden. The comparison with other works is presented in Table 7. First, we will consider the work of those who

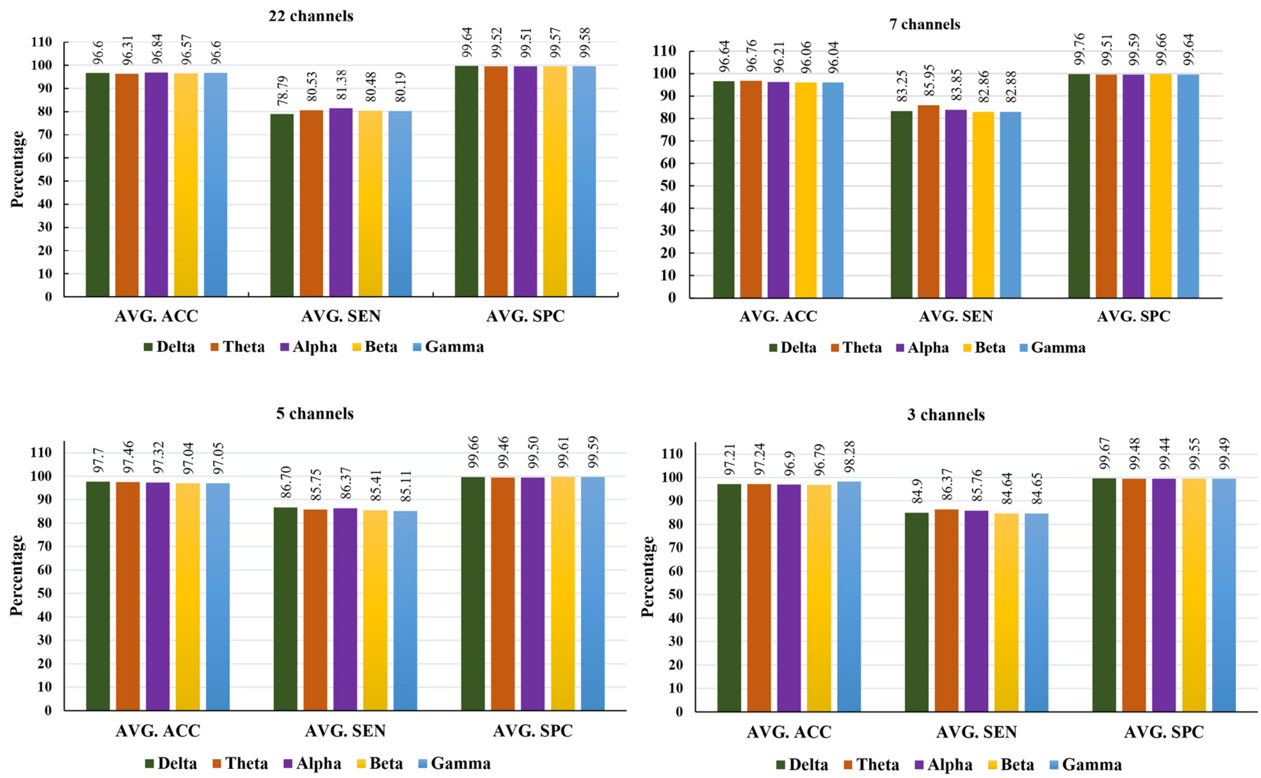

Fig. 8 Average performance metrics across five bands and all patients using the different numbers of channels. **a** original 22 channels, **b** top 7 channels, **c** top 5 channels, and **d** top 3 channels

have utilized the CHB-MIT dataset for implementation purposes. The study [7] has selected the channels after the feature extraction stage, which will not reduce the processing time required in this stage. With the use of 3 channels, authors have achieved a sensitivity of 80.87% whereas, our method has achieved an average sensitivity of 86.70% using 5 channels and the channel has been selected before extracting the feature ultimately saving time in the feature extraction stage. According to [20] the maximum accuracy obtained using 2 channels is 97.5%, which is almost similar to the accuracy obtained by our approach. Accuracy alone may not suffice, as metrics like sensitivity and specificity are crucial for evaluating a model's effectiveness, particularly in critical applications like medical diagnosis. The work [23] has reported a sensitivity of 100% using 8 channels for 6 epileptic patients (chb01, chb02, chb03, chb05, chb08, and chb11) whereas, our approach has obtained an average sensitivity of 85.98% across 24 patients and 98.45% if the same 6 patients are considered using 5 channels. The average sensitivity and accuracy utilizing two channels reported by [2] are 78.9% and 71.91%, respectively, which are significantly less than what our technique has achieved. The results using iEEG electrodes have to be superior to scalp EEG since the original captured signal using iEEG itself contains less noise. The author [12] has obtained 96% of sensitivity utilizing 3iEEG channels which is almost 10% more than ours. Even the study [27] has employed 3iEEG channels. Our proposed ensemble method enhances robustness, reduces biases, enhances the generalization capability of a system, and increases confidence in selected channels. According to our research, using a small subset of relevant EEG channels greatly increases the sensitivity of seizure detection when compared to employing a larger subset of channels. This innovative method emphasizes the possibility of more accurate and efficient seizure detection systems.

Discussion

In this proposed work, we introduced a novel channel selection method using an ensemble of channel rankings obtained from MI-based and RF-based methods. Accuracy, sensitivity, and specificity are standard metrics used to evaluate the performance of a model in seizure detection. Accuracy is an overall performance measure of the seizure detection model. Our proposed method has obtained 97.70% accuracy across all patients and frequency bands suggest the our method

Table 5 Band-wise performance metrics obtained using the original 22 channels and the top 5 channels for patient 1 to patient 13 (all values are in percentage)

Metrics	Original 22 channels					Top 5 channels				
	δ	θ	α	β	γ	δ	θ	α	β	γ
ACC_1	98.61	97.27	96.05	96.05	96.05	99.62	99.00	97.38	97.38	97.38
SEN_1	94.58	87.48	80.67	80.67	80.67	99.42	97.08	87.72	87.72	87.72
SPC_1	99.64	99.79	100	100	100	99.68	99.52	100	100	100
ACC_2	100	100	100	100	100	100	100	100	100	100
SEN_2	100	100	100	100	100	100	100	100	100	100
SPC_2	100	100	100	100	100	100	100	100	100	100
ACC_3	99.41	97.83	99.01	99.01	99.05	99.71	98.84	99.28	99.57	99.42
SEN_3	91.17	92.90	94.09	92.38	92.81	98.77	95.06	96.91	98.15	98.15
SPC_3	99.83	99.28	99.87	99.79	99.70	100	100	100	100	99.81
ACC_4	94.33	95.07	94.39	93.80	93.91	93.47	96.50	94.87	93.47	94.64
SEN_4	83.13	85.61	83.66	80.41	80.51	82.91	87.93	81.90	78.05	83.05
SPC_4	99.09	99.09	98.94	98.94	99.09	97.44	99.68	99.68	99.67	99.04
ACC_5	95.87	98.25	96.89	95.17	94.42	96.87	99.15	98.39	95.52	94.75
SEN_5	72.27	90.96	83.88	75.05	70.59	83.70	96.04	92.95	80.62	76.65
SPC_5	99.77	99.81	96.67	99.49	99.53	100	99.90	99.69	99.06	99.06
ACC_6	94.55	94.90	97.36	98.44	98.75	93.99	95.95	98.50	98.95	98.65
SEN_6	22.13	37.17	62.14	70.25	73.25	57.78	65.00	83.33	88.89	87.04
SPC_6	100	100	99.50	99.81	99.88	99.84	100	99.84	99.84	99.67
ACC_7	98.77	98.55	97.82	98.47	98.99	99.52	99.04	98.44	98.68	99.28
SEN_7	91.35	90.43	85.37	90.24	91.15	96.23	93.28	88.18	90.76	95.08
SPC_7	100	99.94	99.94	99.87	99.71	100	100	100	100	100
ACC_8	97.79	97.56	98.65	98.31	98.16	98.38	99.31	99.15	99.23	99.07
SEN_8	92.70	90.54	94.79	91.74	90.34	93.99	98.23	97.04	97.85	97.03
SPC_8	99.21	99.53	99.73	99.60	99.51	99.80	99.60	99.80	99.61	99.70
ACC_9	99.01	99.79	99.55	99.04	98.97	99.25	99.85	99.4	99.25	99.25
SEN_9	91.09	98.72	98.29	90.15	90.16	94.57	99.05	98.08	98.15	97.06
SPC_9	99.76	100	99.80	99.59	99.47	100	100	99.64	99.46	99.64
ACC_{10}	99.12	99.61	98.94	98.26	98.26	99.60	99.54	99.07	98.07	98.01
SEN_{10}	92.07	99.62	89.72	83.11	83.11	96.61	96.61	92.66	85.31	83.05
SPC_{10}	99.93	99.95	100	100	100	100	99.92	99.92	99.77	100
ACC_{11}	87.68	81.52	81.29	81.54	82.25	99.25	87.86	87.86	86.05	86.69
SEN_{11}	68.90	56.66	52.59	55.08	58.25	98.27	75.92	73.08	65.02	64.95
SPC_{11}	96.67	93.44	95.04	94.22	93.76	99.69	93.44	95.22	96.07	96.45
ACC_{12}	92.66	92.87	92.96	92.96	92.96	93.75	93.50	93.75	92.89	92.59
SEN_{12}	68.00	69.34	69.03	69.03	69.03	71.82	70.22	68.62	69.31	65.53
SPC_{12}	100	100	100	100	100	100	100	100	100	100
ACC_{13}	97.67	97.65	97.67	97.67	97.67	97.46	97.46	97.46	97.46	97.46
SEN_{13}	81.11	85.81	85.81	82.41	82.11	86.41	86.41	86.41	86.41	86.41
SPC_{13}	100	99.97	100	100	100	100	100	100	100	100

ACC_1, SEN_1, and SPC_1: Accuracy, sensitivity, and specificity respectively for patient 1

correctly identifies seizure and non-seizure events. However, relying solely on accuracy in imbalanced datasets can be misleading. High accuracy is desirable, but its importance depends on balancing sensitivity and specificity. Sensitivity is crucial for seizure detection systems where missing a seizure (false negatives) [26] could have serious consequences. Our method has achieved 86.70% of sensitivity, demonstrating good performance in identifying seizures. A model with high sensitivity ensures that most seizure events

Table 6 Band-wise performance metrics obtained using the original 22 channels and the top 5 channels for patient 14 to patient 24 (all values are in percentage)

Metrics	Original 22 channels					Top 5 channels				
	δ	θ	α	β	γ	δ	θ	α	β	γ
ACC_{14}	97.87	97.87	97.87	97.87	97.87	97.74	97.32	98.45	97.88	98.31
SEN_{14}	68.08	68.08	68.08	68.08	68.08	67.35	62.75	73.17	66.67	68.42
SPC_{14}	100	100	100	100	100	100	100	100	100	100
ACC_{15}	86.24	85.94	96.47	92.62	92.77	93.31	95.22	93.17	93.08	92.90
SEN_{15}	60.02	59.66	71.91	78.17	78.47	84.48	88.06	83.55	80.77	79.97
SPC_{15}	98.99	99.14	99.86	99.87	99.96	97.92	98.96	98.2	99.51	99.65
ACC_{16}	99.28	99.28	99.28	99.28	99.28	99.29	99.29	98.70	99.53	99.06
SEN_{16}	70.27	70.37	73.27	70.07	70.13	75.00	68.75	60.71	80.00	75.00
SPC_{16}	100	100	100	100	100	100	99.88	100	100	100
ACC_{17}	96.22	96.47	96.52	96.50	96.45	96.09	96.09	96.43	96.43	96.09
SEN_{17}	68.16	71.91	68.19	69.07	70.65	69.57	70.43	72.17	72.17	72.17
SPC_{17}	100	99.86	99.97	100	100	100	99.87	100	100	99.62
ACC_{18}	98.06	95.58	97.00	95.76	96.47	98.39	96.79	97.27	96.99	96.87
SEN_{18}	89.16	80.36	86.39	80.13	80.12	94.37	88.37	90.99	90.06	89.68
SPC_{18}	99.56	99.34	99.12	98.68	99.34	99.49	99.08	98.98	98.88	98.82
ACC_{19}	100	100	100	100	100	100	100	100	100	100
SEN_{19}	99.14	99.3	99.24	99.21	99.23	100	100	100	100	100
SPC_{19}	100	100	100	100	100	100	100	100	100	100
ACC_{20}	100	100	100	100	100	100	100	100	100	100
SEN_{20}	100	100	100	100	100	100	100	100	100	100
SPC_{20}	100	100	100	100	100	100	100	100	100	100
ACC_{21}	95.29	95.29	95.29	95.52	95.29	96.66	96.66	96.66	96.66	96.66
SEN_{21}	63.31	63.31	63.31	65.09	63.31	69.81	69.81	69.81	69.81	69.81
SPC_{21}	100	100	100	100	100	100	100	100	100	100
ACC_{22}	97.25	98.34	99.63	99.72	99.72	99.08	99.18	100	100	100
SEN_{22}	64.26	79.88	94.50	90.40	90.20	88.71	91.01	100	100	100
SPC_{22}	99.97	99.90	99.97	100	100	99.78	100	100	100	100
ACC_{23}	96.84	96.94	96.82	96.90	96.94	97.34	97.25	96.70	96.98	97.25
SEN_{23}	77.11	78.16	77.71	78.46	73.16	85.63	88.51	85.06	87.36	90.23
SPC_{23}	100	99.95	99.88	99.86	99.95	99.56	98.91	98.91	98.8	98.58
ACC_{24}	95.93	94.76	94.67	94.76	94.15	96.14	95.18	94.86	94.96	94.96
SEN_{24}	83.29	76.45	70.56	72.29	70.77	85.29	84.43	81.22	76.85	75.65
SPC_{24}	99.00	99.39	100	100	100	98.56	98.34	98.14	100	100
$ACC_{AVG.}$	**96.60**	**96.31**	**96.84**	**96.57**	**96.60**	**97.70**	**97.46**	**97.32**	**97.04**	**97.05**
$SEN_{AVG.}$	**78.79**	**80.53**	**81.38**	**80.48**	**80.19**	**86.70**	**85.75**	**86.37**	**85.41**	**85.11**
$SPC_{AVG.}$	**99.64**	**99.52**	**99.51**	**99.57**	**99.58**	**99.66**	**99.46**	**99.50**	**99.61**	**99.59**

The average values of all performance metrics across all patients are presented in bold

ACC_{14}, SEN_{14}, and SPC_{14}: Accuracy, sensitivity, and specificity respectively for patient 14. *AVG.* indicates average values across 24 patients

are captured. Specificity is important for systems where reducing false alarms (false positives) is critical, such as in wearable devices or long-term monitoring. A specificity obtained by our proposed method is of 99.66%, indicating that the method correctly identifies 99.66% of non-seizure events, meaning it generates very few false alarms. High specificity ensures that non-seizure segments are not mistakenly classified as seizures, reducing unnecessary interventions. A good seizure detection system balances sensitivity and specificity according to the needs of the application.

The results obtained by this proposed work as depicted in Fig. 8 show that selecting top 5 channels from the CHB-MIT dataset outperformed the use top 3 and top

Table 7 Comparison with other works

Authors	Methods	Dataset	Channels	Performance measures		
				ACC	SEN	SPC
[12]	Variance	FHS	3 iEEG	–	96	–
[7]	RF	CHB-MIT	3 sEEG	–	80.87	–
[27]	Ensemble classifiers	Kaggle	3 iEEG	–	89.40	89.24
[23]	NCA	CHB-MIT	8 sEEG	–	100	–
[20]	NSGA	CHB-MIT	2 sEEG	97.5	–	–
[13]	Variance	CHB-MIT	3 sEEG	89.02	100	77.5
[2]	NN with attention mechanism	CHB-MIT	2 sEEG	71.91	78.9	–
Proposed work	Ensemble of MI and RF	CHB-MIT	5 sEEG	97.70	86.70	99.66

ACC accuracy, *SEN* sensitivity, *SPC* specificity

7 channels, indicating that a targeted approach to channel selection can provide more relevant information for detecting seizures. The enhancement in sensitivity is prominently seen with reduced channels in almost all patients. An overall improvement of 5% to 8% in sensitivity is observed with the top 5 selected channels. This proposed approach has achieved a mean accuracy of 97.70%, a mean sensitivity of 86.70%, and a mean specificity of 99.66% using five selected channels.

Compared to previous studies that relied on a single ranking method [2, 7, 12, 13], our ensemble-based approach demonstrated superior performance. The combination of MI and RF rankings provided a more comprehensive and accurate selection of channels, leading to better feature representation and classification results. This approach outperformed other standard channel selection criteria, highlighting the importance of integrating multiple ranking methods for optimal channel selection.

In signal processing and neuroscience, band-wise analysis is a powerful technique that enables researchers to focus on particular frequency bands and derive significant information from complex signals. It plays a crucial role in enhancing detection accuracy, understanding physiological processes, and developing targeted solutions for various applications. Certain patterns related to seizures in EEG signals may manifest more prominently in particular frequency bands. By focusing on these bands, analysts can enhance the detection of specific phenomena. Specifically, the δ and θ bands showed a strong performance in identifying seizure occurrences, probably because they have associations with cognitive and motor functions that change dramatically during seizures. In the same way, the γ band which is associated with higher-order cognitive functions has performed admirably, highlighting its importance in capturing the high-frequency changes that are specific to seizures.

As per the proposed methodology, the top 5 channels can be the optimal channels to work with for this dataset. Selecting the top 5 channels strikes an optimal balance between capturing relevant EEG features for accurate seizure detection and maintaining computational efficiency. Using fewer channels risks losing critical information, potentially reducing sensitivity and specificity, while using more channels can lead to redundancy, increased complexity, and overfitting. By limiting the number of channels to 5, we ensure that the model remains efficient, interpretable, and generalizable, while still capturing the essential features needed for accurate seizure detection using the CHB-MIT dataset. This approach strikes a balance between performance and practicality. Hence, the selection of channels is often driven by the trade-off between computational efficiency and the need to capture relevant features for accurate detection, and this balance may change when working with different datasets or research objectives. Hence, the number of channels can be adjusted based on the specific requirements of the study or dataset, allowing flexibility to optimize performance for different scenarios.

Wearable EEG systems have several challenges such as these systems are prone to noise from environmental factors such as movement artifacts and muscle activity. This can degrade the quality of the EEG signals. Wearable devices have constrained processing power and battery life. Despite potential challenges, our proposed method can significantly contribute to wearable devices for seizure detection. MI quantifies the dependency between EEG channels and seizure states to identify the most informative channels, while RF ranks channels based on their contribution to accurate predictions. This process excludes noisy and redundant channels that may carry artifacts from muscle movements, environmental interference, or redundant information to some extent, thus improving the signal-to-noise ratio. Cleaner signals allow for more accurate feature extraction, reducing the impact of noise or irrelevant data on the subsequent classification stage. Even the processing time required using 5 channels is reduced by almost 80% compared to the original 22 channels, as shown in Fig. 6b. Hence reducing the number of channels significantly lowers computational complexity, processing time, and power requirements, which is crucial for wearable devices. This optimization not only conserves energy, but also extends the battery life of the wearable EEG device.

The practical implications of our findings are significant for real-time seizure detection systems. By reducing the number of channels, this proposed method not only enhances detection sensitivity but also lowers the computational load, making the system more suitable for real-time applications. This performance is crucial for wearable devices and mobile health applications, where processing power and battery life are constrained. Although our proposed method offers numerous benefits, it may require advanced noise-reduction techniques for managing the noise from movement, electrode displacement, and environmental interference. Balancing real-time processing, reliable data acquisition, and wireless transmission under strict power constraints poses the challenge for maintaining battery life.

The use of diverse features, combined with an ensemble of MI and RF-based methods, supports a diverse population by effectively capturing patient-specific variations. Together, these methods ensure personalized and reliable seizure detection across varying EEG patterns in a diverse population. Since this method adapts to patterns found in the data, it offers a flexible and adaptive way to identify the essential channels for seizure detection.

Conclusion

This proposed research work provides an effective channel selection methodology, offering a balance between computational efficiency and accurate seizure detection. The proposed channel selection method ensembles the channel ranking obtained by MI-based and RF-based methods providing useful insights into determining the most informative EEG channels for seizure detection. The channel ranking obtained by both methods identifies a comparable set of important channels for seizure detection. MI-based methods excel at capturing explicit dependencies between EEG signals and seizure patterns, making them ideal for identifying relevant channels. In contrast, RF-based methods discover implicit dependencies by evaluating complex interactions and redundancies among channels, ensuring a more comprehensive channel selection process. By combining these methods, we have exploited their complementary strengths to enhance channel selection, leading to more robust and generalized models.

The band-wise analysis, which examines frequency bands such as δ, θ, α, β, and γ, provides a deeper understanding of the spectral characteristics of seizures. This analysis helps identify which frequency bands are most informative for detecting seizures, allowing for more precise channel selection and improved model sensitivity and specificity. Our proposed method has obtained a mean value of accuracy, sensitivity, and specificity of 97.70%, 86.70%, and 99.66% respectively using five selected channels.

Future advancements could focus on optimizing this ensemble approach for real-time applications, integrating deep learning techniques, and data augmentation methods for more enhanced seizure detection. Expanding studies to include diverse patient populations could significantly improve system robustness, ensure broader applicability, and facilitate widespread clinical adoption. Hence, this proposed method can be highly useful in wearable devices, as it helps to optimize performance by reducing computational complexity and processing time while enhancing detection accuracy.

Funding
The authors declare that no funds, grants, or other support were received for this research work.

Data availability
The data used in this work is obtained from publicly available sources: https://physionet.org/content/chbmit/1.0.0/

Declarations

Conflict of interest
The authors have no relevant financial or non-financial interests to disclose.

Author details
[1]Department of Electronics Engineering, K. J. Somaiya College of Engineering, Somaiya Vidyavihar University, Vidyanagar, Vidyavihar East, Mumbai, Maharashtra 400077, India. [2]Department of Computer Engineering, Vivekanand Education Society's Institute of Technology, Chembur, Mumbai 400074, India.

Received: 1 September 2024 Accepted: 3 March 2025
Published online: 19 March 2025

References
1. Acharya UR, Sree SV, Swapna G, Martis RJ, Suri JS. Automated eeg analysis of epilepsy: a review. Knowl-Based Syst. 2013;45:147–65.
2. Affes A, Mdhaffar A, Triki C, Jmaiel M, Freisleben B. Personalized attention-based eeg channel selection for epileptic seizure prediction. Expert Syst Appl. 2022;206:117733.
3. Alotaiby T, El-Samie FEA, Alshebeili SA, Ahmad I. A review of channel selection algorithms for eeg signal processing. EURASIP J Adv Signal Process. 2015;2015:1–21.
4. Berger H. Über das elektroenkephalogramm des menschen. Arch Psychiatr Nervenkr. 1929;87(1):527–70.

5. Bergil E, Bozkurt MR, Oral C. An evaluation of the channel effect on detecting the preictal stage in patients with epilepsy. Clin EEG Neurosci. 2021;52(5):376–85.
6. Biau G, Scornet E. A random forest guided tour. TEST. 2016;25:197–227.
7. Birjandtalab J, Pouyan MB, Cogan D, Nourani M, Harvey J. Automated seizure detection using limited-channel eeg and non-linear dimension reduction. Comput Biol Med. 2017;82:49–58.
8. Chakrabarti S, Swetapadma A, Pattnaik PK. A channel selection method for epileptic eeg signals. In *Emerging technologies in data mining and information security: proceedings of IEMIS 2018*. Vol. 1. Springer; 2019. p. 565–573.
9. Chang N-F, Chen T-C, Chiang C-Y, Chen L-G. Channel selection for epilepsy seizure prediction method based on machine learning. In 2012 Annual international conference of the IEEE Engineering in Medicine and Biology Society; 2012. p. 5162–5.
10. Cortes C, Vapnik V. Support-vector networks Mach Learn. 1995;20:273–97.
11. Cover TM. Elements of information theory. Hoboken: Wiley; 1999.
12. Duun-Henriksen J, Kjaer TW, Madsen RE, Remvig LS, Thomsen CE, Sorensen HBD. Channel selection for automatic seizure detection. Clin Neurophysiol. 2012;123(1):84–92.
13. Ein Shoka AA, Alkinani MH, El-Sherbeny A, El-Sayed A, Dessouky MM. Automated seizure diagnosis system based on feature extraction and channel selection using EEG signals. Brain Informatics. 2021;8(1):1–16.
14. Emara HM, Elwekeil M, Taha TE, El-Fishawy AS, El-Rabaie E-SM, Alotaiby T, Alshebeili SA, Abd El-Samie FE. Hilbert transform and statistical analysis for channel selection and epileptic seizure prediction. Wirel Pers Commun. 2021;116:3371–95.
15. Ghosh-Dastidar S, Adeli H, Dadmehr N. Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. IEEE Trans Biomed Eng. 2007;54(9):1545–51.
16. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. Ccirculation. 2000;101(23):e215–20.
17. Guyon I, Gunn S, Nikravesh M, Zadeh LA. Feature extraction: foundations and applications, vol. 207. Springer; 2008.
18. Iranzad R, Liu X. A review of random forest-based feature selection methods for data science education and applications. Int J Data Sci Anal. 2024. https://doi.org/10.1007/s41060-024-00509-w
19. Khan KA, Shanir P, Khan YU, Farooq O. A hybrid local binary pattern and wavelets based approach for eeg classification for diagnosing epilepsy. Expert Syst Appl. 2020;140: 112895.
20. Moctezuma LA, Molinas M. Eeg channel-selection method for epileptic-seizure classification based on multi-objective optimization. Front Neurosci. 2020;14:593.
21. PhysioNet Chb-mit scalp EEG database. 2010. https://physionet.org/content/chbmit/1.0.0/. Accessed 4 Jan 2022.
22. Qaraqe M, Ismail M, Serpedin E. Band-sensitive seizure onset detection via CSP-enhanced EEG features. Epilepsy Behav. 2015;50:77–87.
23. Sayeed MA, Mohanty S, Kougianos E, Rachakonda L. Rseiz: a channel selection based approach for rapid seizure detection in the IOMT. In 2019 IEEE International symposium on smart electronic systems (iSES) (Formerly iNiS). IEEE; 2019. p. 105–10.
24. Shoeb A, Edwards H, Connolly J, Bourgeois B, Treves ST, Guttag J. Patient-specific seizure onset detection. Epilepsy Behav. 2004;5(4):483–98.
25. Sikdar D, Roy R, Mahadevappa M. Epilepsy and seizure characterisation by multifractal analysis of EEG subbands. Biomed Signal Process Control. 2018;41:264–70.
26. Singh K, Malhotra J. Smart neurocare approach for detection of epileptic seizures using deep learning based temporal analysis of eeg patterns. Multimedia Tools Appl. 2022;81(20):29555–86.
27. Truong ND, Kuhlmann L, Bonyadi MR, Yang J, Faulks A, Kavehei O. Supervised learning in automatic channel selection for epileptic seizure detection. Expert Syst Appl. 2017;86:199–207.
28. Tuncer E, Bolat ED. Channel based epilepsy seizure type detection from electroencephalography (EEG) signals with machine learning techniques. Biocybern Biomed Eng. 2022;42(2):575–95.
29. Wang Z, Mengoni P. Seizure classification with selected frequency bands and EEG montages: a natural language processing approach. Brain Informatics. 2022;9(1):11.
30. WHO. Epilepsy. WHO; 2024. https://www.who.int/news-room/fact-sheets/detail/epilepsy.
31. Yang S, Li B, Zhang Y, Duan M, Liu S, Zhang Y, Feng X, Tan R, Huang L, Zhou F. Selection of features for patient-independent detection of seizure events using scalp EEG signals. Comput Biol Med. 2020;119: 103671.
32. Yildiz M, Bergil E. The investigation of channel selection effects on epileptic analysis of EEG signals. Balkan J Electr Comput Eng. 2015;3:236–41.