

A Novel CNN-FRTAM Model for Enhanced Detection of Epileptic Seizures in EEG Signals Utilizing Deep Learning and Continuous Wavelet Transform

Riyam Ali Yassin¹, Morteza Valizadeh¹, Alaa Hussein Abdulaal^{2,*}, Mehdi Chehl Amirani¹

¹Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran

²Department of Electrical Engineering, College of Engineering, Al-Iraqia University, Baghdad, Iraq

Received 05 February 2025; received in revised form 13 April 2025; accepted 14 April 2025

DOI: <https://doi.org/10.46604/peti.2024.14819>

Abstract

This study presents a novel approach to enhancing electroencephalography (EEG) signal classification for improved epileptic seizure detection. Traditional techniques for seizure detection often rely on manual analysis and suffer from high computational burdens and bias. Integrating deep learning with the continuous wavelet transform (CWT) is proposed to address these limitations for effective feature extraction. The CNN-FRTAM model, which integrates a convolutional neural network (CNN) and a frequency-region temporal attention mechanism (FRTAM), employs rigorous pre-processing of EEG signals to optimize performance. Extensive evaluation on a diverse dataset revealed an accuracy of 99.80% in multi-class classification and 99.90% in binary classification between Normal and Abnormal states. The CNN-FRTAM model significantly outperformed traditional architectures such as InceptionV3, VGG19, and ResNet50, demonstrating its potential for effective real-time applications in clinical epileptic seizure management. By opening up new avenues for accurate seizure detection, this work contributes to improving patient outcomes in epilepsy care.

Keywords: epileptic, EEG signals, CNN, continuous wavelet transform, attention mechanism

1. Introduction

Epilepsy is a prevalent neurological disorder with recurring seizures, affecting approximately 50 million people worldwide [1]. Classification and diagnosis of seizures with high accuracy are critical for effective therapy and management, with an impact on therapeutic interventions and prognosis of the patient. Electroencephalography (EEG) is regarded as a gold standard for brain activity observation and electrical signatures in seizures [2]. Despite this, traditional seizure detection used to, in many cases, require manual analysis, which entails high computational burdens and susceptibility to bias.

The progression of deep learning over the years has changed many fields, including medical signal processing, with its application in the computerized analysis of complex datasets [3]. Techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been effective in processing EEG, leveraging these networks' prowess in mapping high-dimensional information onto complex representations [3]. Despite such advances, most current deep techniques lack an effective integration of rich details in EEG's time and frequency axes.

Attention mechanism techniques have become a powerful tool in deep networks, providing selective processing of important input parts [4]. Such techniques enhance model interpretability and classification accuracy by allowing the

* Corresponding author. E-mail address: EngineerAlaaHusseini@gmail.com

prioritization of important parts of the input [5]. Nonetheless, traditional methods for utilizing attention mechanisms have paid little attention to processing EEG spectrograms in a general 2D format, with the omission of key information in both time and frequency axes, which is important for the effective classification of seizures.

A new model, namely the frequency-region temporal attention mechanism (FRTAM), is proposed to address such vulnerabilities. Specifically for seizure classification in EEG signals, three types of attention mechanisms in FRTAM include frequency attention, for targeting frequency-specific signatures of a seizure; temporal attention, for tracking transient development of a seizure over a period; and region-aware contextual attention, for targeting regions in both time and frequency that are capable of indicating epileptic events.

The proposed FRTAM aims to improve seizure classification accuracy with an interpretable output capable of conveying the driving factors in model decision-making to clinicians. With diligent experiments performed with publicly accessible EEG datasets, the current work aims to test the efficacy of FRTAM in comparison with state-of-the-art techniques, demonstrating its utility in enhancing seizure detection and its contribution to improving the effective management of epilepsy.

2. Related Work

The analysis of EEGs, with their non-stationarity and high temporal fluctuations, requires high-level analysis techniques for effective feature extraction. To address such requirements, Dogra et al. [6] adopted the use of discrete wavelet transform (DWT) for extracting complex frequency contents of EEGs. In its proposed scheme, an optimized k-nearest neighbors (KNN) algorithm was utilized, and surprisingly, accuracy in detections increased, proving effectiveness in feature extraction through waves in seizure detections.

Ryu et al. [7] proposed a new algorithm for seizure detection employing principal component analysis (PCA) for feature extraction. In the algorithm, PCA-based feature extraction was compared with a variety of machine learning classifiers, including logistic regression (LR), dense trees, 2D-support vector machines (2D-SVM), and cosine k-nearest neighbors (cos-KNN). Using PCA for feature reduction, the algorithm improved training and testing performance, supporting feature reduction in model robustness.

The use of deep neural networks, specifically convolutional neural networks (CNNs), in seizure detection continues to improve performance with a 2D model proposed in one study with three layers, processing both time and frequency domains of EEGs. With its use of frequency domain inputs, its performance exhibited an astonishing 95.6% accuracy in seizure detection [8]. Long short-term memory (LSTM) networks, an alternative model of recurrent neural networks (RNNs), have been utilized in numerous studies in their use for seizure detection. One report combining both CNNs and LSTMs exhibited area under the curve (AUC) values of 92% and 84% for Temple University Hospital Seizure Detection Corpus (TUSZ) and Royal Prince Alfred Hospital (RPAH) datasets, respectively [9], proving complementary use of such deep neural architectures in its performance improvement in detections.

The application of Transformers in analyzing EEG signals also gained traction. Ma et al. [10] leveraged a Transformer model with time-frequency representations of EEG, with state-of-the-art performance and an AUC value of 92.1% in the TUSZ dataset. Another study leveraged a self-supervised version of a Transformer for seizure detection, with impressive performance with accuracy and sensitivity values of 88.59% and 89.58%, respectively, proposed by Xiao et al. [11]. One such development involved a new convolutional Transformer model, with 96.71% and 97.23% values for sensitivity and specificity, respectively, proposed by Ke et al. [12], with its efficacy proven in real-life medical use cases.

In addition to conventional classifiers, chi-square tests have been leveraged by Asadi-Pooya et al. [13] to identify fourteen high-correlated features. Then, various classifiers, such as random forests, decision trees, support vector machines, k-nearest neighbors, and TabNet, have been leveraged. In this work, feature selection is emphasized for model performance improvement.

Also, new visualization tools have been leveraged for model decision visualization. Sun et al. [14] showed channel-wise attention values derived through a self-attention mechanism in their model. Zhao et al. [15] leveraged a model attention and gradient-weighted class activation mapping (Grad-CAM) with a residual neural network (ResNet) model for creating heat maps. Yet, such visualization tools lack medical interpretability, and granular visualization of attention maps cannot deliver meaningful information for medical professionals.

To conclude, developments in epileptic seizure detection through analysis of EEG signals have seen considerable growth, and various deep learning methodologies have been leveraged. Exploration of feature extraction, model architectures, and interpretability is critical for developing reliable and applicable solutions in real-life medical environments.

3. Materials and Methods

Epileptic seizures are acute, uncontrollable electrical disturbances of the brain, which produce a variety of symptoms such as convulsions or alterations in awareness. They are either generalized seizures, affecting both hemispheres of the brain, or focal seizures, beginning in a localized region. The typical causes vary, including genetic predisposition, head injury, and infections. Diagnosis is typically through EEG and brain imaging, whereas treatment often involves antiepileptic medications, lifestyle changes, and, in some cases, surgery. With proper management, the majority of individuals with epilepsy lead normal lives.

3.1. Bonn EEG time series dataset

The Bonn EEG database [16] consists of 100 single-channel recordings, each lasting 23.6 seconds, with a 173.61 Hz sampling rate, and 0.5 to 85 Hz bandwidth for its spectra, and these recordings have been acquired with a 128-channel system. The five sets in the database, sets A, B, C, D, and E, represent five subjects, as shown in Fig. 1. Sets A and B represent surface recordings from healthy subjects with eyes open and eyes closed. The sets C and D represent an eyes-open and an eyes-closed, respectively, but not a seizure-related region in an epileptic patient. Set E consists of an intracranial, eyes-open, but seizure-related region in an epileptic patient. The five sets consist of 100 files, each of 4097 samples of a single, in ASCII format, EEG time series. A 0.53 and 40 Hz cutoffs band-pass filtering have been conducted, and, therefore, no preprocessing is conducted, namely, in terms of distinguishing between healthy (non-epileptic) and unhealthy (epileptic) and in filtering out eye-movement-related artifacts and significant eye-movement-related artifacts have been removed. The collection is publicly accessible and first appeared in 2001, and it is a useful source for various studies in the field.

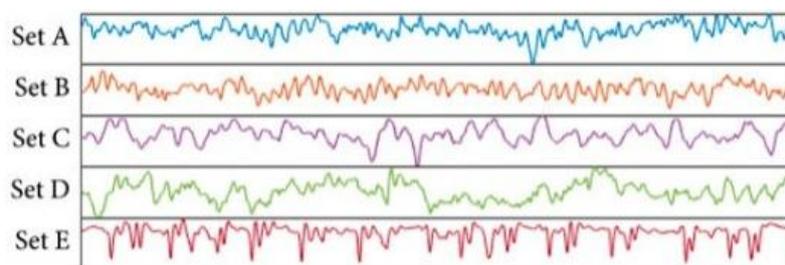


Fig. 1 EEG waveforms in the Bonn dataset [17]

3.2. Epileptic Classification Method

The proposed algorithm for the classification of epileptic signals transforms traditional one-dimensional (1D) EEG signals into two-dimensional (2D) EEG images in a format capable of leveraging strong convolutional neural networks (CNNs) for image analysis, as shown in Fig. 2. The raw EEG signals initially undergo preprocessing for artifact and rejection of noise, as well as for integrity in terms of data. The 1D signals then become represented in 2D format in spectrograms for visualization and feature extraction. To simplify generalization, various data augmentation techniques are leveraged in a format to expand

diversity in the dataset. A frequency-region temporal attention mechanism (FRTAM) is added to the model in a format that enables selective region focusing in 2D images and salience improvement for classification. Lastly, processed images undergo classification with a complex CNN model, with the added use of an attention mechanism for accuracy.

The proposed in-depth study aims to significantly enhance the accuracy of epileptic classification by coupling state-of-the-art signal processing techniques with deep learning methodologies. By harnessing the strengths of both approaches, this project claims to create a more reliable and efficient paradigm in epilepsy diagnosis and treatment. This sort of combined system has the potential to improve patient outcomes, facilitate early intervention, and provide healthcare professionals with valuable insights into seizure patterns and causes. Ultimately, this study aims to clear the path for more effective treatment methods and a better quality of life for patients with epilepsy.

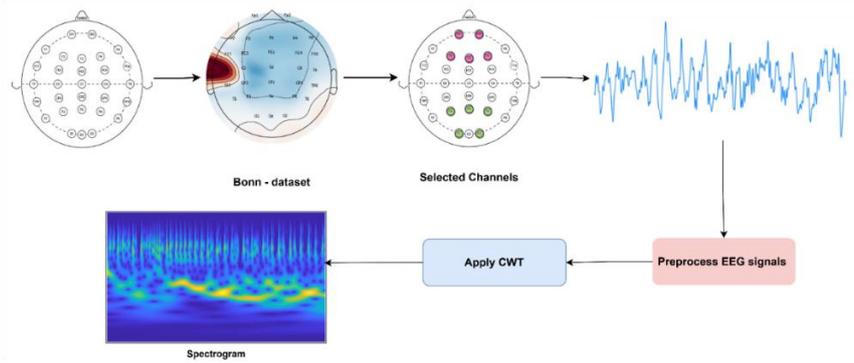


Fig. 2 Signal Converting process

3.3. Signal Preprocessing

Effective signal processing is paramount for proper classification and correct diagnosis of a seizure in an EEG record. In this part, significant preprocessing operations in the processing of an EEG signal, culminating in the construction of 2D spectrograms, the continuous wavelet transform (CWT) of the Bonn dataset, serve as a starting point for use with the proposed frequency-region temporal attention mechanism (FRTAM).

3.3.1. Preprocessing

Preprocessing of EEGs forms an important stage for any subsequent analysis to produce reliable and correct output. This section discusses numerous preprocessing techniques that have been utilized to improve the quality of EEGs.

(1) Filtering

EEG signals often contain noise and artifacts that can obscure meaningful brain activity. To address this, a bandpass filter allows only the frequency components of interest to pass through [18]. The filtering process can be mathematically represented by the following convolution operation

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau \tag{1}$$

where $x(t)$ is the raw EEG signal, $h(t)$ is the filter's impulse response, and $y(t)$ is the filtered output signal.

The impulse response $h(t)$ can be synthesized for a bandpass filter Eq. (2) using the difference of two low-pass Eq. (3) and high-pass filters Eq. (4). For example, using a Butterworth filter, the transfer function for the bandpass filter can be defined as

$$H(f) = \frac{H_{high}(f) \cdot H_{low}(f)}{H_{high}(f) + H_{low}(f)} \tag{2}$$

$$H_{low}(f) = \frac{1}{1 + \left(\frac{f}{f_c^{low}}\right)^{2n}} \quad (3)$$

$$H_{high}(f) = \frac{1}{1 + \left(\frac{f_c^{high}}{f}\right)^{2n}} \quad (4)$$

where f is the frequency, f_c^{low} , and f_c^{high} are the cutoff frequencies for the low-pass and high-pass filters, respectively.

n is the filter order. This filter design attenuates frequencies outside the specified range (e.g., 1-40 Hz).

(2) Normalization

Normalization is crucial to standardize the amplitude of EEG signals across different recordings. The normalization procedure Eq. (5) can be defined as follows

$$x_{norm}(t) = \frac{x(t) - \mu}{\sigma} \quad (5)$$

where μ is the mean of the signal, and can be obtained by

$$\mu = \frac{1}{N} \sum_{i=1}^N x(i) \quad (6)$$

and σ is the standard deviation, and can be calculated by

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x(i) - \mu)^2} \quad (7)$$

This transformation ensures that the normalized signal has a mean of 0 and a standard deviation of 1, mitigating variations due to electrode placement and individual differences.

(3) Artifact Removal

EEG signals frequently become contaminated with many artifacts, including eye and muscle, and external electrical activity. One of the most common techniques for decomposing such an artifact and isolating it from actual brain activity is independent component analysis (ICA), which can be calculated by

$$x = As \quad (8)$$

where x is the observed EEG signal, A is the mixing matrix, and s is the vector of independent source signals, which can be found by Eq. (9)

The goal is to estimate A and s such that the components in s are statistically independent. This is achieved by maximizing the non-Gaussianity of the estimated components, often using measures such as kurtosis or mutual information.

The separation can be achieved using algorithms like Fast ICA, which iteratively updates the unmixing matrix W

$$s = Wx \quad (9)$$

where W is updated based on the maximization of the negentropy, which is described as

$$J(W) = H(y) - H(x) \tag{10}$$

where H denotes entropy.

(4) Segmentation

The continuous EEG signal is segmented into epochs to analyze specific interest segments, typically ranging from 1 to 10 seconds. This segmentation can be represented as

$$x_{epoch}[j] = x[t_j : t_{j+1}] \tag{11}$$

where t_j and t_{j+1} define the start and end times of the j -th epoch.

Epoching helps isolate seizure events from the continuous EEG and ensures that the analysis focuses on relevant timeframes.

(5) Detrending

EEG signals can exhibit non-stationary characteristics, leading to trends that interfere with analysis. Detrending is performed to remove these trends as described by the below linear regression model

$$x_{detrended}(t) = x(t) - (a + bt) \tag{12}$$

where a and b are the coefficients obtained from fitting a line to the data over the epoch. This step ensures that the analysis focuses on fluctuations around a mean value, enhancing the detection of seizure-related patterns.

3.3.2. Spectrogram Generation

The transformation of pre-processed EEG signals into a time-frequency format is important for properly classifying seizures [19]. In the following section, creating spectrograms by applying the continuous wavelet transform (CWT) brings in a lengthy mathematical formalism for the overall scheme.

(1) continuous wavelet transform (CWT)

The continuous wavelet transform analyzes the signal at various scales, allowing for time and frequency localization [20]. The CWT of a signal $x(t)$ is defined as

$$CWT(a,b) = \int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t-b}{a}\right)dt \tag{13}$$

where $\psi(t)$ is the wavelet function, a is the scale parameter, which inversely relates to frequency, and b is the translation parameter (time shift).

The wavelet function $\psi(t)$ must satisfy certain properties, such as being localized in time and frequency. A commonly used wavelet function is the Morlet wavelet, defined as

$$\psi(t) = \frac{1}{\sqrt{2\pi}} \left(e^{i\omega_0 t} e^{-\frac{t^2}{2}} \right) \tag{14}$$

where ω_0 is the central frequency of the wavelet. The choice of ω_0 influences the frequency resolution of the transform. The scale a is related to frequency f as follows

$$f = \frac{\omega_0}{a} \quad (15)$$

(2) Calculation of CWT Coefficients

For each scale a and translation b , the CWT provides a complex coefficient $CWT(a,b)$. The magnitude of these coefficients is computed to represent the strength of the signal at different scales and times

$$|CWT(a,b)| = \sqrt{Re(CWT(a,b))^2 + Im(CWT(a,b))^2} \quad (16)$$

where Re and Im denote the real and imaginary parts of the CWT coefficient, respectively.

This magnitude represents how much of a particular frequency (scale) is present in the signal at a given time.

(3) Time-Frequency Representation

The resulting CWT coefficients can be arranged into a 2D matrix, where the x-axis represents time b , the y-axis represents scale a (or equivalently frequency), and the values represent the magnitudes of the CWT coefficients. This matrix forms the basis of the time-frequency representation, often visualized as a spectrogram.

(4) Frequency Scaling

To convert the scale a back to frequency f , the following relationship is used

$$f = \frac{f_0}{a} \quad (17)$$

where f_0 is the reference frequency corresponding to the wavelet, this conversion is essential for interpreting the results regarding natural frequency bands (delta, theta, alpha, beta, gamma).

(5) Spectrogram Construction

The spectrogram $S(t, f)$ can be constructed from the magnitude of the CWT coefficients

$$S(t, f) = |CWT(a,b)| \quad \text{for each } (a,b) \quad (18)$$

where $S(t,f)$ is the spectrogram, t represents time, and f represents frequency. The computed spectrogram provides a detailed view of how the frequency content of the EEG signal evolves.

(6) Resizing and Normalization of the Spectrogram

To ensure uniformity across the dataset, the generated spectrograms are resized to a fixed dimension ($T \times F$), where T is the number of time steps, and F is the frequency bins. To resize, it can be performed using interpolation techniques, such as bilinear interpolation, defined as

$$S'(x, y) = \sum_{i=0}^1 \sum_{j=0}^1 S(i, j) \cdot w(i, x) \cdot w(j, y) \quad (19)$$

where $w(i,x)$ and $w(j,y)$ are the interpolation weights based on the distance to the neighboring points.

In addition to resizing, normalization of the spectrogram can be performed to enhance contrast and improve the effectiveness of subsequent classification models, that is

$$S_{norm}(t, f) = \frac{S(t, f) - S_{min}}{S_{max} - S_{min}} \quad (20)$$

where S_{min} and S_{max} are the minimum and maximum values in the spectrogram, respectively. This normalization ensures that the values fall within a range of [0, 1].

The spectrogram generated via CWT of processed EEGs creates a rich and powerful expression of information regarding time-frequency behavior, which is important for the efficient classification of seizures. Such a technique can extract individual character transience and individual contents regarding frequency for a specific type of seizure. With such particular definitions in mathematics, the preprocessing and transformation of EEGs become powerful. They can effectively act to drive complex deep neural networks such as the frequency-region temporal attention mechanism (FRTAM).

3.4. Pre-trained Deep Convolutional Neural Networks (DCNN)

Pre-trained deep convolutional neural networks (DCNNs) have revolutionized computer vision and have been applied in numerous cases, including seizure classification for spectrograms of EEG. In this section, the three most popular architectures, VGG19, ResNet50, and InceptionV3, will be discussed briefly.

3.4.1. VGG19

VGG19 is a 19-layer deep trainable-parameter network with a uniform structure. VGG19 is distinguished by a uniform configuration with 3×3 small convolutional filters in each network section [21]. With this configuration, VGG19 can extract fine-grained spatial information effectively and with a relatively low number of parameters. VGG19's deep structure is its most salient feature, through which hierarchical feature representations can be obtained. VGG19 has been proven to perform well in most image classification problems and is utilized regularly in feature extraction in cases of transfer learning. It is, therefore, applicable for processing spectrograms.

3.4.2. ResNet50

ResNet50 introduces skip connections with residual learning, enabling deep networks to be trained without the vanishing gradient problem [22]. With 50 layers, ResNet50 promotes efficient gradient flow through layers, improving training convergence in return. Composed of residual blocks, it allows simple identity mappings to be learned easily, with complex mappings becoming simple for a model to learn. ResNet50 performs best in local and global feature extraction, a feature beneficial in scenarios with rich information analysis, such as in an EEG spectrogram for seizure detection.

3.4.3. InceptionV3

InceptionV3 is an efficient architecture with mixed-dimensional filters in one layer, and with its use, multi-scale feature extraction can effectively be attained [23]. In its development, sometimes an "Inception module" can learn about numerous regions of the input information in a single pass, and with its performance and efficiency, InceptionV3 can use fewer parameters compared with traditional architectures, while still achieving high accuracy. Methods such as batch normalization and auxiliary classifiers are part of its training and performance improvement architecture. InceptionV3's feature extraction can make it a powerful tool for processing EEG spectrograms, in which complex relationships in both dimensions of time and frequency become significant for the efficient classification of seizures.

4. Proposed Model

The proposed model consists of a frequency-region temporal attention mechanism (FRTAM) and convolutional neural networks (CNNs) to enhance the epileptic signal analysis. FRTAM is used to find important frequency regions and temporal

patterns in the data, allowing the model to focus on important features for improved classification accuracy. By integrating this attention mechanism with the robust feature extraction strength of CNNs, the model strives to give a more accurate and effective method to diagnose epilepsy and ultimately improve patient management and treatment plans.

4.1. Frequency-Region Temporal Attention Mechanism (FRTAM)

The frequency-region temporal attention mechanism (FRTAM) is designed to interpret 2D spectrograms effectively with a view toward three important aspects in epileptic events: localized regions and development in both dimensions over a while, and structures specific in terms of frequency. In a coherent and organized manner, it enables neural networks to classify seizures precisely.

4.1.1. Frequency Attention

Epileptic seizures are associated with distinct frequency bands. For example, delta waves (0.5-4 Hz) are prevalent in certain seizures, while fast oscillations, such as gamma waves (30-100 Hz), may indicate different seizure types, which are described as

$$S \in \mathbb{R}^{T \times F} \quad (21)$$

Eq. (21) denote the 2D spectrogram, where T is the number of time steps, and F is the number of frequency bins.

A learnable weight vector, $w_f \in \mathbb{R}$, is introduced to compute attention scores for each frequency bin, capturing the importance of different frequency bands during seizure events. The attention weights α_f are calculated as follows

$$\alpha_f = \text{softmax}(w_f \cdot S^T) \in \mathbb{R}^F \quad (22)$$

where the dot product $w_f \cdot S^T$ yields a score for each frequency bin, and the SoftMax function converts these scores into a probability distribution, which can be ensured that

$$\sum_{f=1}^F \alpha_f = 1 \quad (23)$$

This normalization indicates the relative importance of each frequency bin. The spectrogram is weighted along the frequency axis as shown in

$$S_{\text{focused}} = S \odot \alpha_f \quad (24)$$

In Eq. (24), the symbol \odot denotes element-wise multiplication. This operation dynamically prioritizes frequency bands relevant to seizure detection, allowing the model to emphasize critical features.

4.1.2. Temporal Attention

Seizures evolve and exhibit distinct phases, including onset, maintenance, and termination. Temporal attention captures this evolution by modeling dependencies across the time axis. A bidirectional gated recurrent unit (BiGRU) processes the frequency-focused spectrogram as shown in

$$h_t = \text{BiGRU}(S_{\text{focused}}) \quad (25)$$

where h_t represents the hidden states at each time step, encoding the temporal dynamics of the spectrogram. Attention weights for each time step are computed using a learnable query vector w_t

$$\alpha_t = \text{softmax}(w_t \cdot h_t) \quad (26)$$

This calculation generates weights that reflect the importance of each time step in the seizure evolution. The temporally attended feature representation is obtained by applying the attention weights to the hidden states, as shown in

$$S_{temporal} = \sum_{t=1}^T \alpha_t \cdot h_t \quad (27)$$

This summation ensures that the model focuses on critical time intervals, such as seizure onset or significant activity bursts, while disregarding irrelevant background signals.

4.1.3. Region-Based Contextual Attention

Spectrograms encode joint time-frequency information, and local regions often reveal patterns crucial for distinguishing between seizure and non-seizure states. The spectrogram is divided into non-overlapping regions, or patches, typically of size (e.g. 8×8). A lightweight CNN extracts features from each patch, as shown in

$$P = \text{CNN}(S) \quad (28)$$

where P denotes the feature representation of the patches.

Self-attention weights for each patch are computed based on their similarity to a global learnable embedding vector. This is represented as

$$\alpha_r = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right) \quad (29)$$

where Q is Query matrix, K is Key matrix, and d is Dimension of the key vectors, This equation uses scaled dot-product attention, ensuring that the resulting weights α_r are normalized. The region-focused spectrogram is generated by weighting the patch features, which are described as

$$S_{region} = \sum_r \alpha_r \cdot V \quad (30)$$

where V represents the value matrix, this operation allows the model to focus on local time-frequency regions indicative of epileptic events.

4.1.4. Integration of All Attention Mechanisms

The outputs from the three attention mechanisms are combined to form a comprehensive feature representation, which is described by

$$S_{combined} = [S_{focused}, S_{temporal}, S_{region}] \quad (31)$$

This combined representation integrates the insights gained from frequency, temporal, and region-based attention mechanisms.

4.1.5. Weighted Feature Representation

To further enhance the discriminative power of the combined features, a transformation layer is applied

$$Z = W \cdot S_{combined} + b \quad (32)$$

where Z is the transformed feature representation, W is a learnable weight matrix, and b is a bias vector. This transformation enables the model to learn an optimal representation that maximizes the separation between seizure and non-seizure classes.

4.1.6. Classification Layer

The processed features are passed through a fully connected layer to produce the final classification output, as shown in

$$Y = \text{softmax}(W_c \cdot Z + bc) \quad (33)$$

where Y represents the predicted probability distribution over classes (e.g., seizure vs. non-seizure), W_c represents the weight matrix for the classification layer, and b represents the bias term.

4.1.7. Loss Function

To train the model, the categorical cross-entropy loss function is employed:

$$L = -\sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{P}(y_{i,c})) \quad (34)$$

where N is the number of samples, C is the number of classes, $y_{i,c}$ is the true label for sample i and class c , and $\hat{P}(y_{i,c})$ is the predicted probability for sample i and class c . This loss function quantifies the difference between the predicted and actual class distributions, guiding the model's training process.

4.2. Integration of CNN with FRTAM

Integrating deep convolutional neural networks (DCNNs) with the frequency-region temporal attention mechanism (FRTAM) is a robust platform for processing EEG spectrograms for seizure classification. This integration takes advantage of feature extraction with DCNNs and taps into the interpretability and performance capabilities of FRTAM. The pipeline begins by utilizing the first feature extractor in a DCNN to process the input spectrograms. The DCNN learns to extract salient structures in the spectrograms. The hierarchical feature in spectrograms is captured effectively through several convolution layers in a DCNN, generating rich feature maps that encode salient information in the input, as shown in Table 1. and Fig. 3. The produced feature maps then go through processing in the FRTAM module, in which three forms of attention mechanism are employed: Frequency Attention, Temporal Attention, and Region-Based Contextual Attention.

Frequency Attention identifies important frequency bands concerning a seizure, with model prioritization of important spectral information. Temporal attention identifies important phases in developing a seizure over time, with model prioritization of important phases, such as initiating a seizure. Region-based contextual Attention identifies important regions in localized time frequencies, with model prioritization of important structures indicative of a seizure. The output of these three forms of attention mechanism is then merged into a single feature representation.

Merging is achieved through simple concatenation, generating a single, merged feature representation combining strong feature extraction of a DCNN with focused model attention of an FRTAM. A transformation layer is added to further enhance such a merged feature map. The transformation layer utilizes a learnable bias vector and a weight matrix to maximize feature suitability for classification. The feature map is used in a fully connected layer, and a classification output is generated, providing a probability distribution over classes.

Incorporating DCNN with FRTAM, as shown in Fig. 4, helps enhance model accuracy and incorporates an attention mechanism and interpretability. Information about the most important frequency bands, times, and regions involved in predicting a seizure is uncovered through an attention mechanism. Overall, such a composite model constitutes a significant

contribution to developing epileptic seizure classifiers, providing a strong, efficient, and transparent system for processing EEG. Table 1 describes DCNN integration with the incorporation of FRTAM layers, with the direction of information and purpose of each constituent in the overall structure.

Table 1 2D-DCNN model layers

Layer	Type	Kernel Size	Stride	Kernel	Input Size
1-2	2D Convs	3×3	1	64	$224 \times 224 \times 3$
3	Pooling	2×2	2	-	$112 \times 112 \times 64$
4-5	2D Convs	3×3	1	128	$112 \times 112 \times 64$
6	Pooling	2×2	2	-	$56 \times 56 \times 128$
7-9	2D Convs	3×3	1	256	$56 \times 56 \times 128$
10	Pooling	2×2	2	-	$28 \times 28 \times 256$
11-13	2D Convs	3×3	1	512	$28 \times 28 \times 256$
14	Pooling	2×2	2	-	$14 \times 14 \times 512$
15-17	FRTAM Attention Mechanisms	-	-	-	$14 \times 14 \times 512$
18	Concatenation	-	-	-	$14 \times 14 \times 1536$
19	Fully Connected Layer 1	-	-	4096	$14 \times 14 \times 1536$
20	Fully Connected Layer 2	-	-	3	4096
21	Output	-	-	-	3

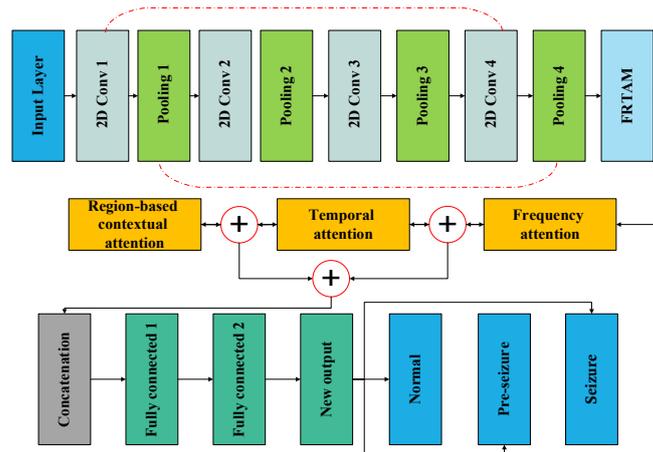


Fig. 3 The architecture of the 2D-DCNN model

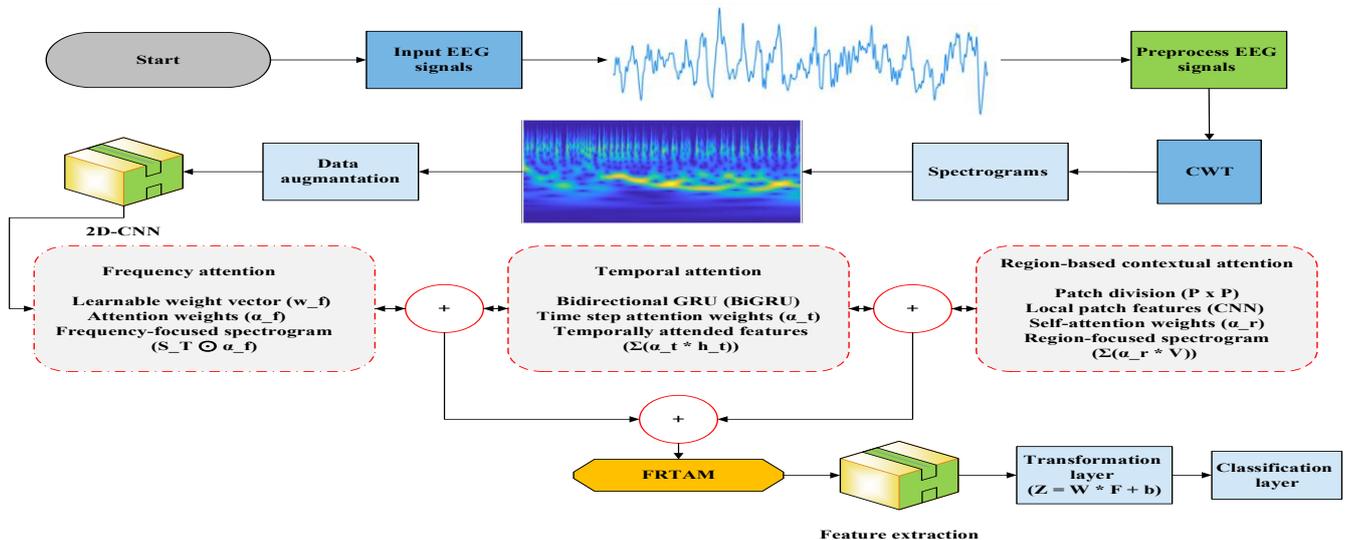


Fig. 4 Flow chart of the proposed model

4.3. SoftMax Activation Function

A SoftMax activation function is important for mapping model output to readable probabilities for multi-class classification [24]. SoftMax transforms model output logits and normalizes them to a sum-to-one probability. The model output can then be understood as a probability for a certain state or kind of seizure [24]. The SoftMax function is defined as:

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (35)$$

where z_i represents the logit for class i , and the denominator sums the exponentials of all logits, ensuring that the outputs are normalized. The advantages of using SoftMax in FRTAM include enhanced decision-making and sensitivity to logit differences, which help the model focus on the most likely seizure type, thereby improving diagnostic accuracy. The SoftMax activation function is pivotal for transforming model outputs into actionable probabilities, facilitating effective multi-class classification, and supporting medical decision-making processes related to seizure disorders.

4.4. Cross-validation

4-fold cross-validation is a systematic method for testing model performance. In 4-fold cross-validation, a dataset is partitioned randomly into four balanced sets (folds). In one run, one fold is retained for testing, and three are used for training a model. It is performed four times, with one fold used for testing in each run. Performance in accuracy, precision, recall, and F1-score for a fold being held out is measured following model training. The values from all four runs for such performance estimates are averaged to evaluate model performance fully. 4-fold cross-validation curtails overfitting, maximizes information utilization, and generates a reliable estimate of model generalizability, and therefore a sound conclusion regarding its efficacy in diagnosing seizures.

4.5. Evaluation metrics

The evaluation metrics used in this paper are as follows [23]:

Accuracy: measures the proportion of correctly classified instances out of the total cases.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (36)$$

Recall (also known as Sensitivity) assesses the model's ability to identify all relevant instances.

$$Recall = \frac{TP}{TP + FN} \times 100 \quad (37)$$

Precision: quantifies the accuracy of the positive predictions made by the model.

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (38)$$

Specificity: measures the proportion of actual negatives that are correctly identified.

$$Specificity = \frac{TN}{TN + FP} \times 100 \quad (39)$$

F1 score: The $F1$ score is the harmonic mean of precision and recall, balancing the two. It is particularly useful in cases where class distribution is imbalanced.

$$F1 - Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \times 100 \tag{40}$$

The receiver operating characteristic (ROC) curve represents the model's performance across various thresholds. It plots the true positive rate (Recall) against the false positive rate (1-Specificity). The area under the ROC curve (AUC) provides a single measure of overall model performance, with values closer to 1 indicating better performance. The definitions are as follows: TP: True Positives, FP: False Positives, FN: False Negatives, and TN: True Negatives

5. Results

The results section presents experimental observations of the work in classifying EEG signals for predicting a seizure. The performance of the proposed model, CNN-FRTAM, was evaluated in terms of accuracy, sensitivity, specificity, and F1 score. In Table 2, Multi-class performance reveals that the proposed model, CNN-FRTAM, can accurately classify Normal, Pre-seizure, and Seizure states. The model achieved an overall accuracy of 99.80%, outperforming traditional architectures such as InceptionV3, VGG19, and ResNet50. The model's high accuracy reflects its performance in predicting seizure events for all classes.

Table 2 Evaluation metrics of multi-class classification

CNN	Metrics	Accuracy	Precision	F1-Score	Specificity
InceptionV3	Normal	92.00%	90.20%	91.09%	95.00%
	Pre-seizure	92.00%	93.88%	92.93%	97.00%
	Seizure	94.00%	94.00%	94.00%	97.00%
	Average	92.67%	92.69%	92.67%	96.33%
VGG19	Normal	94.00%	92.08%	93.03%	95.96%
	Pre-seizure	93.88%	93.99%	93.94%	97.00%
	Seizure	92.00%	93.84%	92.91%	96.98%
	Average	93.29%	93.30%	93.29%	96.65%
ResNet50	Normal	94.00%	95.92%	94.95%	98.00%
	Pre-seizure	94.00%	94.00%	94.00%	97.00%
	Seizure	96.00%	94.12%	95.05%	97.00%
	Average	94.67%	94.68%	94.67%	97.33%
CNN-FRTAM	Normal	100.00%	99.40%	99.70%	99.70%
	Pre-seizure	99.40%	100.00%	99.70%	100.00%
	Seizure	100.00%	100.00%	100.00%	100.00%
	Average	99.80%	99.80%	99.80%	99.90%

In the binary classification scenario, as shown in Table 3, the model CNN-FRTAM's accuracy in distinguishing between Abnormal (Seizure) and Normal states was 99.90%, a high accuracy representing its efficacy in diagnosing a seizure attack with fewer incorrect diagnoses. High accuracy reflects that the model effectively separates normal brain activity from a seizure attack, and its real-time application in a clinic for diagnosing seizure attacks is highly probable.

In Fig. 5, a 3x3 grid of deep dream representations exhibits explicit features elicited in the FRTAM layer, distinguishing between EEG signal categories: Normal, Pre-seizure, and Seizure. Individual subplots represent explicit visualization patterns for the model considering frequency profiles, temporal fluctuations, and localized regions. The images present complex processing of EEG information via a neural network, providing a deeper analysis of its classification mechanism. Overall, the visualization expresses complex operations of an attention mechanism in seizure detection.

Table 3 Evaluation metrics for Binary classification

CNN	Metrics	Accuracy	Precision	F1-Score	Specificity
InceptionV3	Normal	92.00%	93.88%	92.93%	94.00%
	Abnormal	94.00%	92.16%	93.07%	92.00%
	Average	93.00%	93.02%	93.00%	93.00%
VGG19	Normal	90.00%	91.84%	90.91%	92.00%
	Abnormal	92.00%	90.20%	91.09%	90.00%
	Average	91.00%	91.02%	91.00%	91.00%
ResNet50	Normal	90.00%	93.75%	91.84%	94.00%
	Abnormal	94.00%	90.38%	92.16%	90.00%
	Average	92.00%	92.07%	92.00%	92.00%
CNN+FRTAM	Normal	99.80%	100.00%	99.90%	100.00%
	Abnormal	100.00%	99.80%	99.90%	99.80%
	Average	99.90%	99.90%	99.90%	99.90%

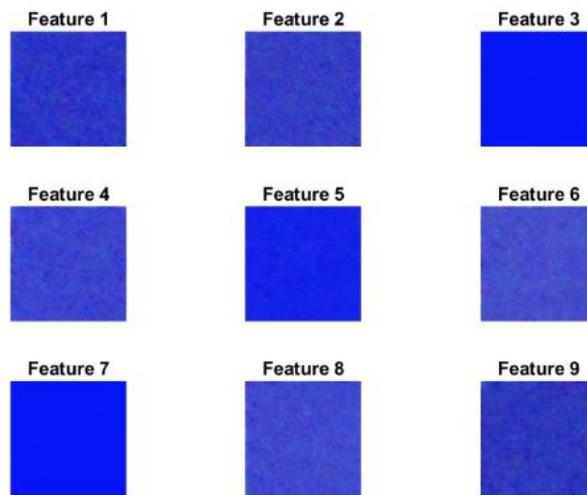


Fig. 5 Deep dream with FRTAM

In Fig. 6, feature extraction presents information about a multi-class classification neural network trained to classify Normal, Pre-seizure, and Seizure classes for EEG signals. In all plots, feature activations for a specific test image are represented as a bar plot. The X-axis represents feature indices, and the y-axis represents activation values. It is a feature contribution visualization that illustrates the model’s prediction for a specific class.

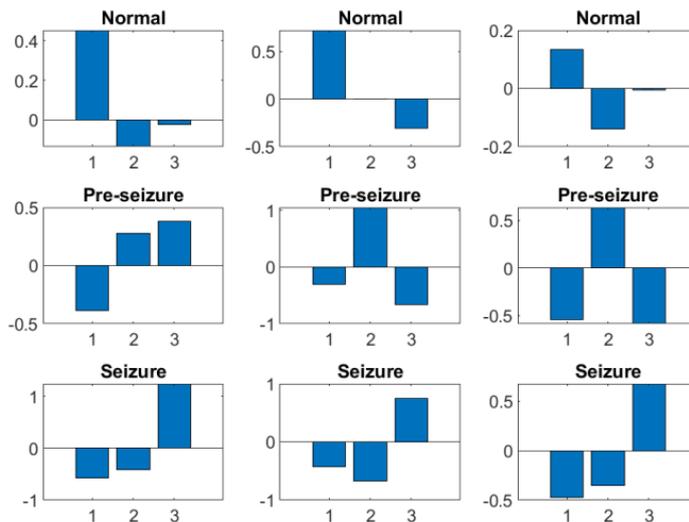


Fig. 6 Features visualization of the last layer

In Fig. 7, visualization via an attention map portrays an enriched view of a neural network's attention toward regions in an EEG spectrogram during its prediction. Each map identifies regions with a high contribution to model prediction for the Normal, Pre-seizure, and Seizure phases. Warm colors denote high attention, and the map identifies significant areas of time and frequency that a model considers significant in differentiation between classes. Providing interpretability, unveiling a model's ability to pay attention to important features while processing complex EEG signals.

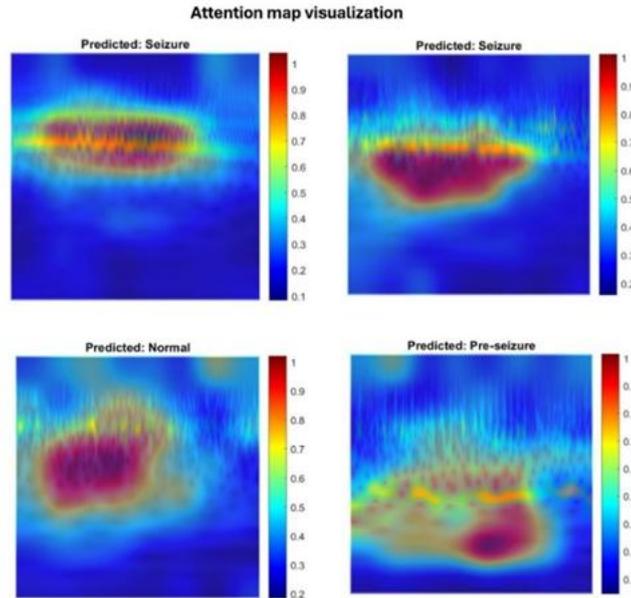


Fig. 7 Attention map visualization

Figs. 8 and 9 illustrate model performance in terms of classification via confusion matrices. Fig. 6 is a multi-class confusion matrix for EEG signals (Normal, Pre-seizure, Seizure) with normalized correct and incorrect prediction count values. Fig. 7 is for binary classification between Seizure and Non-seizure, used to determine model sensitivity and specificity. Together, they illustrate an overall model performance evaluation regarding classification for various scenarios.

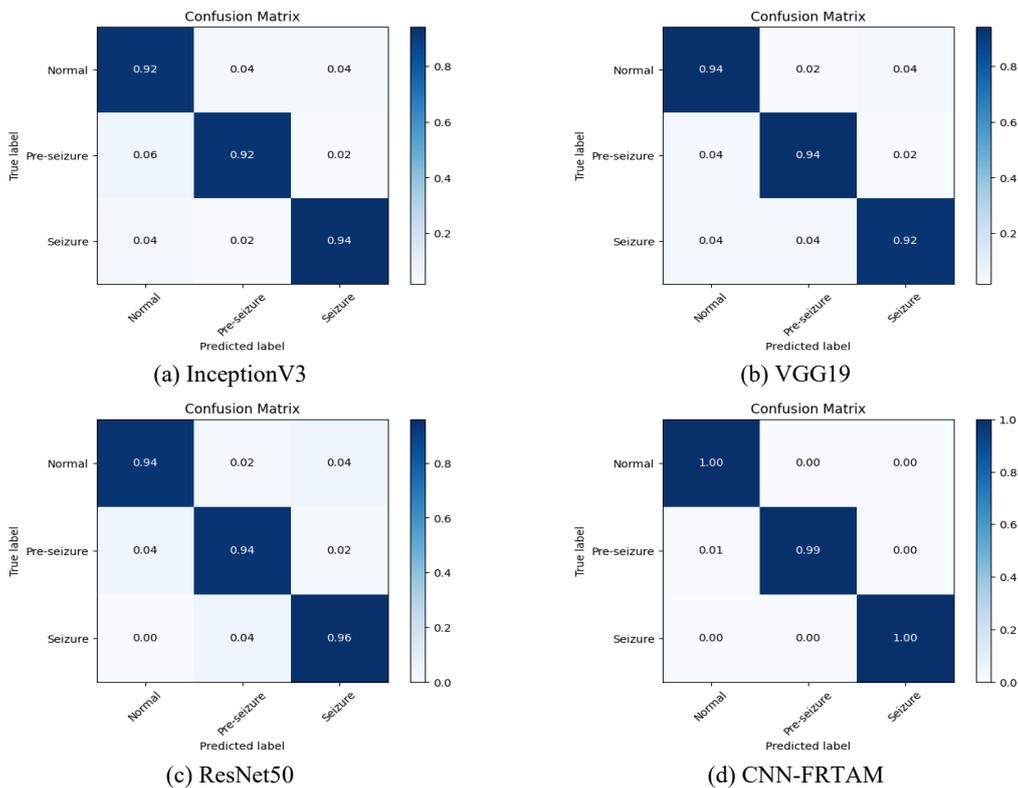


Fig. 8 Confusion matrices (multi-class)

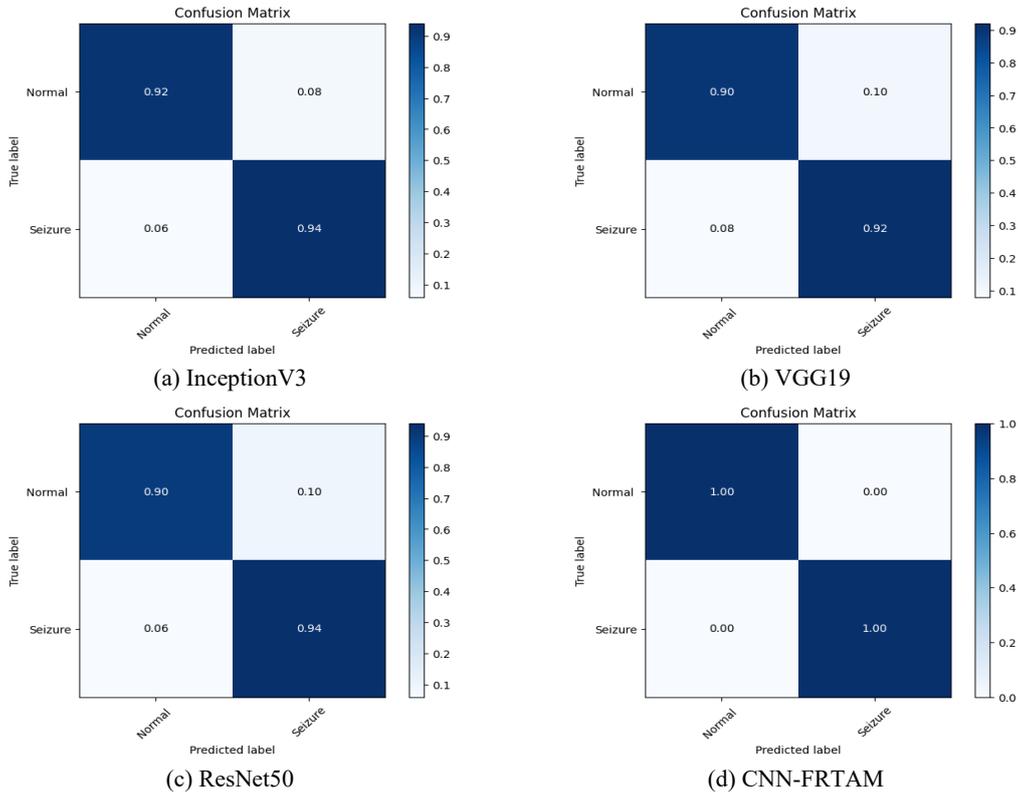


Fig. 9 Confusion matrices (Binary)

The precision-recall and ROC plots are significant performance measurement tools for a model in a classification problem. Precision-Recall plot indicates how recall and precision change with changing decision thresholds, Fig. 10. On the other hand, the ROC plot shows the trade-offs between the true positive and false positive at changing thresholds, Fig. 11. Both individual plots present information about a model distinguishing between classes and choosing an appropriate threshold for a classification problem.

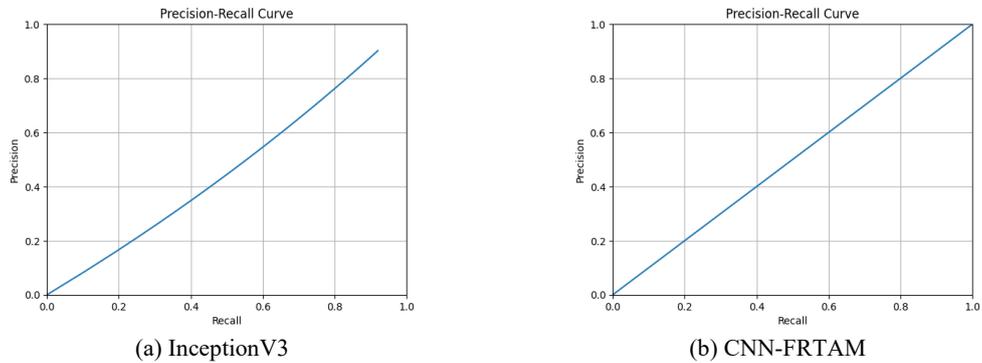


Fig. 10 The Precision-Recall curve

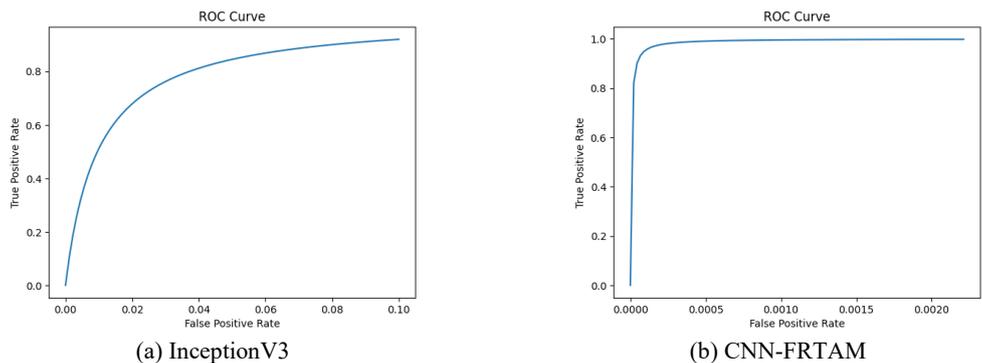


Fig. 11 The ROC curve

Fig. 12 shows the accuracy and loss curves for the proposed CNN+FRTAM model. Table 4 summarizes significant works in EEG signal classification, methodologies, and performance evaluation, and portrays a general picture of development in the field, trends, and gaps in current work to inform future development and research in seizure detection and analysis of EEG.

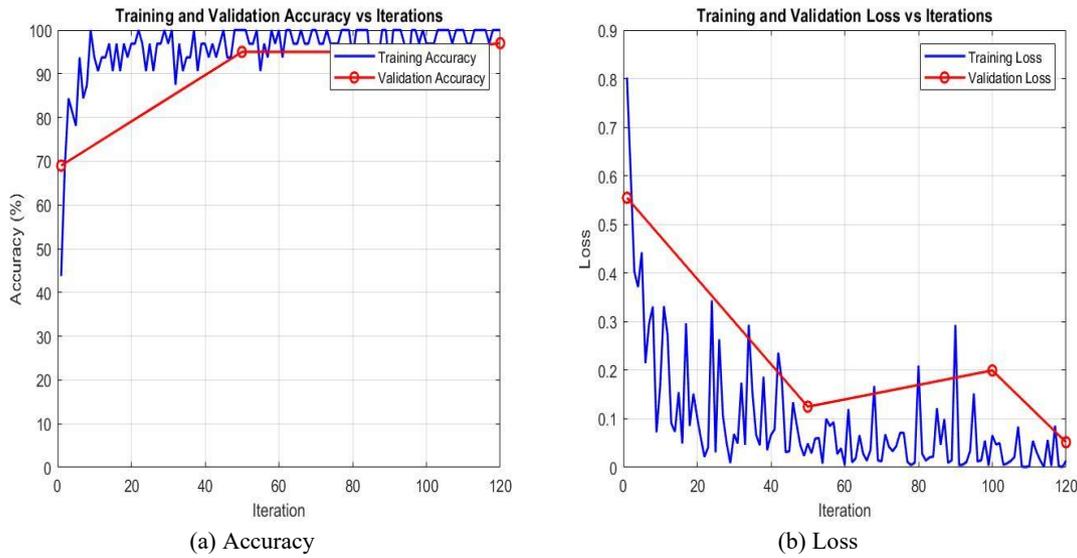


Fig. 12 The accuracy-loss curve for CNN-FRTAM

The attention mechanisms in the FRTAM model should enhance the interpretability since it assigns weights between different segments of EEG signals based on features most indicative of seizures, directing clinicians to the most important parts of the data. In principle, this will create a better ground for informed decisions. This will enable personalized monitoring strategies and treatment plans and, hence, better patient care and improved communication with the patient and their families.

This study acknowledges the limitations of the CNN-FRTAM model. The potential biases in the training data may be a limitation of the study, especially if there is an underrepresentation of some seizure types or demographics, which may affect the sensitivity to the positive performance of the model. Therefore, it may not perform well in noisy environments or with atypical seizure presentations, leading to false alarms. Also, even when trying to ensure interpretability, some of the resultant rule sets are very complex and, therefore, may hinder clinical application. Recognizing these limitations is essential for a comprehensive understanding of the model.

Table 4 Comparison with other related works

Years	Model	Accuracy %	Specificity %	Sensitivity %	Precision %	F1 score
2021	Bi-LSTM [25]	98.79%	-	98.72%	98.86%	98.79%
2022	Random Forest TabNet [13]	64.80%	-	68.25%	64.10%	-
		70.36%	-	80.95%	68.12%	-
2022	CNN Bi-LSTM [26]	93.90%	-	-	-	-
		97.20%	-	-	-	-
2023	CNN-Transformer [27]	82.00%	62.00%	89.00%	-	-
2023	CNN+Bi-LSTM [28]	99.41%	-	98.99%	-	-
2025	DCNN-Transformer-MLP [29]	85.00%	85.00%	82.00%	-	-
2025	This Work CNN+FRTAM	99.80%	99.90%	-	99.80%	99.80%
		99.90%	99.90%	-	99.90%	99.90%

6. Conclusions

This study introduced a novel CNN-FRTAM model that significantly enhances the classification of EEG signals for epileptic seizure detection. By integrating the frequency-region temporal attention mechanism (FRTAM) and employing continuous wavelet transform (CWT) for effective feature extraction, the model demonstrated exceptional performance, achieving an accuracy of 99.80% in multi-class classification and 99.90% in binary classification between normal and abnormal states. The results indicate that the CNN-FRTAM model surpasses traditional architectures like InceptionV3, VGG19, and ResNet50 and holds considerable promise for clinical applications. Its ability to accurately detect seizures in real time conduces to better management and treatment options for individuals with epilepsy.

Future work will refine and test the model on more diverse datasets to ensure robustness and generalizability across different populations and seizure types. Advancements in EEG signal classification will contribute to the broader goal of improving patient outcomes in epilepsy management.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] J. Zhang, S. Zheng, W. Chen, G. Du, Q. Fu, and H. Jiang, "A Scheme Combining Feature Fusion and Hybrid Deep Learning Models for Epileptic Seizure Detection and Prediction," *Scientific Reports*, vol. 14, article no. 16916, 2024.
- [2] V. V. Grubov, S. I. Nazarikov, S. A. Kurkin, N. P. Utyashev, D. A. Andrikov, O. E. Karpov, et al., "Two-Stage Approach with Combination of Outlier Detection Method and Deep Learning Enhances Automatic Epileptic Seizure Detection," *IEEE Access*, vol. 12, pp. 122168–122182, 2024.
- [3] A. H. Abdulwahhab, A. H. Abdulaal, A. H. T. Al-Ghraiiri, A. A. Mohammed, and M. Valizadeh, "Detection of Epileptic Seizure Using EEG Signals Analysis Based on Deep Learning Techniques," *Chaos, Solitons & fractals*, vol. 181, article no. 114700, 2024.
- [4] A. H. Abdulaal, R. A. Yassin, M. Valizadeh, A. H. Abdulwahhab, A. M. Jasim, A. J. Mohammed, et al., "Cutting-Edge CNN Approaches for Breast Histopathological Classification: the Impact of Spatial Attention Mechanisms," *ShodhAI: Journal of Artificial Intelligence*, vol. 1, no. 1, pp. 109–130, 2024.
- [5] Y. Tang, Q. Wu, H. Mao, and L. Guo, "Epileptic Seizure Detection Based on Path Signature and Bi-LSTM Network with Attention Mechanism," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 304–313, 2024.
- [6] A. Dogra, S. A. Dhondiyal, and D. S. Rana, "Epilepsy Seizure Detection Using Optimised KNN Algorithm Based on EEG," *Proceedings of 2023 International Conference for Advancement in Technology*, pp. 1–6, 2023.
- [7] S. Ryu, S. Back, S. Lee, H. Seo, C. Park, K. Lee, et al., "Pilot Study of a Single-Channel EEG Seizure Detection Algorithm Using Machine Learning," *Child's Nervous System*, vol. 37, no. 7, pp. 2239–2244, 2021.
- [8] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, et al., "Epileptic Seizure Detection Based on EEG Signals and CNN," *Frontiers in Neuroinformatics*, vol. 12, article no. 95, 2018.
- [9] Y. Yang, N. D. Truong, C. Maher, A. Nikpour, and O. Kavehei, "Continental Generalization of a Human-In-The-Loop AI System for Clinical Seizure Recognition," *Expert Systems with Applications*, vol. 207, article no. 118083, 2022.
- [10] Y. Ma, C. Liu, M. S. Ma, Y. Yang, N. D. Truong, K. Kothur, et al., "TSD: Transformers for Seizure Detection," <https://www.biorxiv.org/content/10.1101/2023.01.24.525308v2>, 2023.
- [11] T. Xiao, Z. Wang, Y. Zhang, Hongbin lv, S. Wang, H. Feng, et al., "Self-supervised Learning with Attention Mechanism for EEG-Based Seizure Detection," *Biomedical Signal Processing and Control*, vol. 87, article no. 105464, 2024.
- [12] N. Ke, T. Lin, Z. Lin, X. H. Zhou, and T. Ji, "Convolutional Transformer Networks for Epileptic Seizure Detection," *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 4109–4113, 2022.
- [13] A. A. Asadi-Pooya, M. Kashkooli, A. Asadi-Pooya, M. Malekpour, and A. Jafari, "Machine Learning Applications to Differentiate Comorbid Functional Seizures and Epilepsy from Pure Functional Seizures," *Journal of Psychosomatic Research*, vol. 153, article no. 110703, 2022.
- [14] Y. Sun, W. Jin, X. Si, X. Zhang, J. Cao, L. Wang, et al., "Continuous Seizure Detection Based on Transformer and Long-Term iEEG," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 11, pp. 5418–5427, 2022.

- [15] X. Zhao, N. Yoshida, T. Ueda, H. Sugano, and T. Tanaka, "Epileptic Seizure Detection by Using Interpretable Machine Learning Models," *Journal of Neural Engineering*, vol. 20, no. 1, article no. 015002, 2023.
- [16] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of Nonlinear Deterministic and Finite-Dimensional Structures in Time Series of Brain Electrical Activity: Dependence on Recording Region and Brain State," *Physical Review E*, vol. 64, article no. 061907, 2002.
- [17] F. Hassan, S. F. Hussain, and S. M. Qaisar, "Epileptic Seizure Detection Using a Hybrid 1D CNN-Machine Learning Approach from EEG Data," *Journal of Healthcare Engineering*, vol. 2022, no. 1, article no. 9579422, 2022.
- [18] Y. Badr, U. Tariq, F. Al-Shargie, F. Babiloni, F. Al Mughairbi, and H. Al-Nashash, "A Review on Evaluating Mental Stress by Deep Learning Using EEG Signals," *Neural Computing and Applications*, vol. 36, no. 21, pp. 12629–12654, 2024.
- [19] M. Z. Ullah and D. Yu, "Grid-Tuned Ensemble Models for 2D Spectrogram-Based Autism Classification," *Biomedical Signal Processing and Control*, vol. 93, article no. 106151, 2024.
- [20] C. J. O. Echeverri, S. Salazar-Colores, and G. Hernández-Nava, "A Study of the Relationship of Wavelet Transform Parameters and Their Impact on EEG Classification Performance," *Brain-Computer Interfaces*, pp. 115–130, 2025.
- [21] A. H. Abdulaal, M. Valizadeh, R. A. Yassin, B. M. Albaker, A. H. Abdulwahhab, M. C. Amirani, et al., "Unsupervised Histopathological Sub-Image Analysis for Breast Cancer Diagnosis Using Variational Autoencoders, Clustering, and Supervised Learning," *Journal of Engineering and Sustainable Development*, vol. 28, no. 6, pp. 729–744, 2024.
- [22] R. Anand, S. V. Lakshmi, D. Pandey, and B. K. Pandey, "An Enhanced ResNet-50 Deep Learning Model for Arrhythmia Detection Using Electrocardiogram Biomedical Indicators," *Evolving Systems*, vol. 15, pp. 83–97, 2023.
- [23] A. H. Abdulaal, M. Valizadeh, M. C. Amirani, and A. F. M. S. Shah, "A self-learning Deep Neural Network for Classification of Breast Histopathological Images," *Biomedical Signal Processing and Control*, vol. 87, article no. 105418, 2024.
- [24] Y. Fu, C. Zhou, T. Huang, E. Han, Y. He, and H. Jiao, "SoftAct: a High-Precision Softmax Architecture for Transformers Supporting Nonlinear Functions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 9, pp. 8912–8923, 2024.
- [25] A. Abdelhameed and M. Bayoumi, "A Deep Learning Approach for Automatic Seizure Detection in Children with Epilepsy," *Frontiers in Computational Neuroscience*, vol. 15, 2021.
- [26] S. M. Beeraka, A. Kumar, M. Sameer, S. Ghosh, and B. Gupta, "Accuracy Enhancement of Epileptic Seizure Detection: a Deep Learning Approach with Hardware Realization of STFT," *Circuits, Systems, and Signal Processing*, vol. 41, pp. 461–484, 2022.
- [27] W. Y. Peh, P. Thangavel, Y. Yao, J. Thomas, Y. L. Tan, and J. Dauwels, "Six-Center Assessment of CNN-Transformer with Belief Matching Loss for Patient-Independent Seizure Detection in EEG," *International Journal of Neural Systems*, vol. 33, no. 03, 2023.
- [28] I. Ahmad, X. Wang, D. Javeed, P. Kumar, O. W. Samuel, and S. Chen, "A Hybrid Deep Learning Approach for Epileptic Seizure Detection in EEG Signals," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–12, 2023.
- [29] S. Wong, A. Simmons, J. Rivera-Villicana, S. Barnett, S. Sivathamboo, P. Perucca, et al., "Channel-Annotated Deep Learning for Enhanced Interpretability in EEG-Based Seizure Detection," *Biomedical Signal Processing and Control*, vol. 103, article no. 107484, 2025.



Copyright© by the authors. Licensee TAETI, Taiwan. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC) license (<http://creativecommons.org/licenses/by/4.0/>).