I.J. Intelligent Systems and Applications, **2026, 4, 183-204**
Published Online on August 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2026.04.10

Implementing a Novel Framework for Focal and Generalized Epilepsy Classification using Adaptive VAE with Dense Bi-GRU through Multimodal Data-guided Feature Fusion

Maneesh Kumar*
MMMUT, Gorakhpur, Uttar Pradesh, India
E-mail: maneeshniet1987@gmail.com
ORCID iD: https://orcid.org/0009-0003-8112-2914
*Corresponding author

Rakesh Kumar
MMMUT, Gorakhpur, Uttar Pradesh, India
E-mail: rkiitr@gmail.com
ORCID iD: https://orcid.org/0000-0003-1125-0331

Santosh Kumar
Health Economics Unit, NHS, England, UK
E-mail: mrsonuk@gmail.com
ORCID iD: https://orcid.org/0000-0002-6556-4224

Received: January 25, 2026; Revised: March 2, 2026; Accepted: April 3, 2026; Published: August 8, 2026

Abstract: Epilepsy is a neurological condition that affects the emotional and psychological well-being of individuals. Managing this disorder is challenging, particularly because focal seizures, which begin in specific brain regions, can sometimes evolve into generalized forms. The fundamental method for seizure identification is the analysis of the electroencephalogram (EEG), yet its manual interpretation is prone to error. In addition, the process of automated seizure detection using EEG data suffers from variations between subjects and dataset distributions as they might result in inconsistencies. Moreover, the unequal proportion of seizure to non-seizure samples increases the detection challenge. Hence, developing classification approaches capable of distinguishing and predicting the focal and generalized seizures is important for effective treatment planning. Therefore, an efficient deep learning-based focal and generalized epilepsy classification is designed in this research by considering the multimodal data. Initially, essential signals used for the validation are sourced from publicly available EEG datasets and they are converted into Short-Time Fourier Transform (STFT) images, which are designated as Feature Set 1. Next, the Sensor data used for the validation are gathered from Kaggle (https://www.kaggle.com/datasets/datasetengineer/epilepsy-dataset) and it is designated as Feature Set 2. Next, the two acquired sets of features are input to the Multilevel Spatio-Temporal Attention Fusion Network (MSTAFN) to execute the feature fusion process. Once the feature fusion procedure is completed, the fused features are given as the input to the focal and generalized epilepsy classification phase. In this phase, Adaptive Variational Autoencoders with Dense Bidirectional Gated Recurrent Units (AVDBiGRU) are employed to perform the classification process. Moreover, the focal and generalized epilepsy classification process is improved by optimizing the hyperparameters of AVDBiGRU through Fitness-based Football Optimization Algorithm (FFbOA). Finally, the focal and generalized epilepsy classified outcome is obtained from AVDBiGRU. Further, various experiments are carried out in the developed focal and generalized epilepsy classification model over the widely adopted deep learning architectures like LSTM, DCNN, InceptionV3 and BiGRU to verify their efficiency over different classes. The proposed model is evaluated on an EEG dataset containing the high-frequency oscillation (HFO) annotations from 30 pediatric patients with epilepsy and model performance is assessed by using the standard evaluation metrics that includes accuracy, sensitivity, specificity, and F1-score. The proposed model achieved 95.54% accuracy, 96.66% specificity, 93.32% F1-score and 88.53 AUC.

Index Terms: Focal and Generalized Epilepsy Classification, Short-Time Fourier Transform, Multilevel Spatio-Temporal Attention Fusion Network, Adaptive Variational Autoencoders with Dense Bidirectional Gated Recurrent Unit, Fitness-based Football Optimization Algorithm.

1. Introduction

Epilepsy is a neurological condition in which the normal electrical signals of the brain are disrupted, leading to sudden and unpredictable seizures. These episodes can appear in the form of sudden changes in emotions, muscle stiffness, and shaking [9]. Depending on the origin of abnormal brain activity, epilepsy is categorized into two types. In focal epilepsy, the seizures begin in a specific part of the brain [10]. Generalized epilepsy involves seizures that engage both hemispheres of the brain from the onset. Although medications are effective in reducing or preventing seizures, they do not work for every individual [11]. Around one-third of patients still suffer from seizures even after using Anti-Seizure Drugs (ASD). For these individuals, the doctors perform surgery to control the seizures by recognizing the exact brain region that triggers the seizures. The starting point of the seizure activity region is known as the Seizure Onset Zone (SOZ) [12]. After identifying these regions, the doctors might remove the SOZ to reduce seizures. Analyzing the cause of focal and generalized seizures is one of the main challenges in epilepsy research [13]. In many cases, focal epilepsy involves complex brain networks rather than being confined to a single isolated area. Furthermore, focal seizures are capable of triggering secondarily generalized absence seizures. Absence epilepsy is a type of generalized epilepsy characterized by brief lapses in consciousness [14].

A major challenge in epilepsy diagnosis is processing long-term EEG recordings. In order to capture the electrical activity produced by brain neurons through the scalp, EEG sensors are used [15]. EEG analysis is considered as the main technique for monitoring and diagnosing epilepsy. Neurologists identify seizures by manually inspecting EEG data, a process prone to inconsistencies [16]. The time requirement for the diagnostic process is reduced by the development of automated detection technologies. In EEG recording, the electrodes are placed on the scalp surface or probes inserted inside the skull to collect the details regarding the brain activity [17]. Automated EEG-based methods are capable of improving the precision and speed of epilepsy detection [18]. Identifying the most meaningful EEG features for precise seizure classification requires various advanced feature extraction and classification techniques [19].

EEG wavelet spectra resemble image-like structures, allowing seizure detection to be formulated as an image classification problem [20]. Deep learning techniques are capable of learning useful patterns from EEG data for epileptic seizure detection. However, the accuracy and stability of these models are often low when dealing with limited samples [21]. Some studies focus on epilepsy detection using a small number of samples. These approaches contribute to the development of more personalized treatment plans [22]. Existing detection techniques are divided into two main groups: deep learning-based and feature-based methods [23]. Deep learning-based methods like Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs) are data-driven and they automatically learn useful features from raw EEG signals [24]. Due to their ability to learn features directly from data without manual intervention, deep learning models often outperform feature-based techniques. Feature-based methods require manual extraction of signal characteristics for identifying epileptic seizures [25]. Therefore, a focal and generalized epilepsy classification model using deep learning is introduced in this research.

The major contributions of this study are as follows:

- Establishment of a multimodal deep learning framework for focal and generalized epilepsy classification. This research introduces a multimodal diagnostic approach that integrates EEG signals with sensor-based physiological information for analyzing the epileptic activity. The integration of these data sources enhances the ability of the developed model to handle variability across patients and seizure conditions. Distinguishing between focal and generalized epilepsy facilitates appropriate treatment planning.
- Design of an MSTAFN for multimodal feature fusion of STFT-derived spectrogram images and sensor data. The MSTAFN module fuses spatial and temporal patterns after processing multi-level feature domains. This fusion strategy effectively integrates the most informative, context-aware, and semantically rich representations, thereby enhancing classification performance. The non-stationary nature of brain signals is precisely captured by the spatio-temporal attention mechanism as it minimizes the inter-subject variability issues.
- Development of the AVDBiGRU network for identifying focal and generalized epilepsy. The VAE and DBiGRU components process contextual correlations from the fused features to mitigate overfitting. The VAE mechanism enhances the generalizability, while the dense skip connections improve the gradient flow and prevent information loss across the network. Complex temporal features from EEG are precisely handled by the AVDBiGRU model.
- Proposal of an FFbOA to enhance the learning efficiency of the AVDBiGRU network. This algorithm is utilized to fine-tune the hyperparameters of AVDBiGRU to improve the epilepsy identification performance. This optimization overcomes local minima issues via the fitness-based random variable update and it helps to

attain faster convergence. The detection precision and accuracy are enhanced by optimizing hidden neuron count and learning rate.

A focal and generalized epilepsy classification model using a deep learning technique by analyzing multi-modal data is presented in this work. Section II describes the existing epilepsy detection works. Section III showcases the architecture of the designed classification model and the details of the optimization approach. The transformation of input signals into spectrograms and the feature fusion procedures are explained in Sub-class IV. The classification phase is elaborated in Sub-class V. The experimental evaluation and the conclusions are presented in Sub-class VI and Sub-class VII.

2. Literature Survey

2.1. Related Works

In 2022, Najafi et al. [1] developed an epilepsy classification approach by integrating Discrete Wavelet Transform (DWT) and Longitudinal Bipolar montage (LB). At first, EEG signals were decomposed, and relevant statistical features were selected. The extracted data were fed into an LSTM for the classification of epilepsy. Their approach categorized normal, focal, and generalized epilepsy signals using optimized feature selection and deep learning mechanisms.

In 2024, Fan et al. [2] introduced a three-node motif reduction model to represent healthy, focal and critical brain regions for analyzing seizure propagation in focal epilepsy. The authors defined a novel seizure propagation marker to characterize the different seizure patterns and transitions. The influence of critical node excitability and connection heterogeneity on the seizure evolution was proved by the experimental findings.

In 2024, Gill et al. [3] presented a seizure classification model by analyzing real-time multi-class EEG data. They retrieved statistical and spectral features from the input. A hybrid network with the combination of CNN and multi-head self-attention was employed to categorize the seizure types. The complexity of the EEG was effectively handled by the hybrid approach.

In 2022, Narin [4] implemented a focal and non-focal epilepsy prediction model utilizing two-dimensional CNN architectures and Continuous WT (CWT). EEG signals were transformed into 2D scalogram images and classified using multiple pre-trained CNN models. As per the findings, the InceptionV3 technique provided the best performance. Therefore, the suggested method was useful in assisting neurologists for rapid seizure detection.

In 2025, Mokhiamar et al. [5] designed a hybrid model by integrating CNN and LSTM architectures for EEG-based seizure detection. Additionally, the researchers considered blockchain for secure data management. The spectrogram images captured the temporal and spatial patterns, while the hybrid technique improved the learning efficiency. The model showed better detection accuracy and it ensured data privacy through blockchain integration.

In 2025, Hong et al. [6] proposed a self-supervised attention for seizure detection. A hybrid temporal–spatial learning strategy was adopted in this work along with multi-scale feature extraction. They considered a self-attention module based on electrode adjacency. Through extensive clinical dataset evaluations, the network achieved consistent improvements over conventional methods.

In 2025, Li et al. [7] developed an epilepsy classification model by combining the CNN and Vision Transformer (ViT) using a multi-stream feature fusion approach. Initially, the EEG signals were transformed, and the results were processed by the CNN-ViT approaches to capture the local and global patterns. The discriminative ability of the executed model was high due to its feature fusion technique.

In 2025, Yi [8] introduced a deep learning-aided pulse neural epilepsy detection technique with recurrent spiking structures. Their model attained efficient seizure detection results while lowering power consumption. So, the implemented model was suitable for wearable and low-power medical devices.

2.2. Problem Statement

An epileptic seizure is a neurological event, which is characterized by abnormal neuronal discharges. Automatic classification of focal and generalized epilepsy is important to tackle several complications experienced by individuals. Currently, various deep learning techniques utilizing EEG signals have been developed to achieve better outcomes. Yet, these techniques need enormous training samples to process the network. However, their efficiency is often limited by the scarcity of training samples. Several issues that arise while designing a focal and generalized epilepsy classification framework will be listed as follows.

- Traditional classification techniques for focal and generalized epilepsy often yield suboptimal results due to noise and low-amplitude signals. Poor generalization issues in the network lead to inaccurate outcomes and biasing issues.
- Noise in the samples can lead to misdiagnosis, and also take more time to process the samples in the training phase. Moreover, maintaining the generalizability is complex due to poor understanding rates.
- Maintaining the stability of the network is difficult, and the implementation expense is also higher due to the use of enormous parameters. In some cases, the network requires extensive manual verification and consumes

significant computational time.
- Overfitting issues that arise in the deep learning networks need to be tackled, and also require improving the network's robustness in complex classes. Variations in the samples often require more time for validation.
- Reducing the false positive rate is complicated to improve the overall network performance. Moreover, the parameters need to be set properly for collecting the spectral and temporal patterns.

On considering these research gaps, an advanced deep learning-based focal and generalized epilepsy classification framework is developed in this work. Advancements and complications presented in the classical focal and generalized epilepsy classification models are provided in Table 1.

Table 1. Features and challenges of classical focal and generalized epilepsy classification models

Author [citation]	Methodology	Features	Challenges
Najafi *et al.* [1]	LSTM	It recognizes the subtle, evolving seizure patterns over time. It minimizes the need for manual feature engineering and provides better classification accuracy.	Its training time is slow and needs a lot of data for better training. It is sensitive to noise and is also computationally intensive.
Fan *et al.* [2]	Three-node motif reduced network	It minimizes the complex brain regions into 3 simplified nodes. It recognizes the different patterns of seizure spread.	It may lead to loss of nuances and intricate details. It has poor real-world applicability and needs robust validation.
Gill *et al.* [3]	Deep CNN	It automatically earns the features from input, preventing the need for complex feature engineering. It demonstrated improved efficiency and high classification accuracy.	Its training process can be computationally complex. Its interpretability is low and susceptible to overfitting.
Narin [4]	InceptionV3	It minimizes the computational burden and retrieves the nonlinear features. It can generate high-dimensional feature vectors.	It need a high amount of labelled training data. It is sensitive to artefacts and noise.
Mokhiamar *et al.* [5]	LSTM-CNN	It is good in tackling the gradient vanishing issues. It is good in attaining better outcomes by processing the sequential information in different ranges.	Its implementation expense is higher. It has a complicated architecture and requires more samples for training.
Hong *et al.* [6]	MCAN	It has higher interpretability and also maintains robustness in complex classes.	Its outcomes are highly dependent on the input quality. It always use high-quality information.
Li *et al.* [7]	CNN-ViT	It quickly collects the global and local features. Discriminative efficiency of the system is improved with higher robustness.	It takes an enormous time to collect the long-range dependencies. Its implementation is expensive.
Yi [8]	ANN	It reduces the energy utilization and uses limited resources for the validation. It processes the samples in parallel form.	It needs to tackle the overfitting issues. Handling the categorical data in complex.

3. Implementation of Focal and Generalized Epilepsy Classification Model Based on Multimodal Data Using Deep Learning

3.1. Architecture of Proposed Epilepsy Classification Framework

This work proposes an automatic and effective deep learning-based identification method for focal and generalized epilepsy. In this model, multimodal data are considered to improve the accuracy of the epilepsy classification task. The proposed model is helpful in recognizing the focal and generalized epilepsy by analyzing the subtle and localized changes in brain activity. In this model, EEG and sensor data are collected for recognizing focal and generalized epilepsy. Brain activity patterns are effectively monitored by the proposed technique as it processes the temporal and spatial features of EEG as well as sensor data. The input EEG signals are converted into STFT spectrogram images to efficiently convey the information in the form of temporal and frequency domains. The dynamic shifts in brain wave patterns are preserved in the spectrogram images, which is essential for identifying epileptic activity. These images are designated as Feature Set 1. Sensor data from public databases are designated as Feature Set 2. More contextual information is obtained from the sensor data that helps to strengthen the epilepsy classification process. An MSTAFN is employed to extract spatial and temporal features from EEG and sensor data, thereby integrating multimodal information. The multilevel STA model is beneficial in reducing noise. This multilevel fusion strategy results in more discriminative and informative fused features. The suggested MSTAFN improves the epilepsy classification accuracy due to its capacity to handle heterogeneous data sources while maintaining the inherent relationships between the data. An AVDBiGRU network is proposed for the focal and generalized epilepsy classification. This model processes the fused features obtained from the MSTAFN. The VAE model in this network is capable of learning robust latent representations while decreasing the dimensionality of the fused features. The DBiGRU model is responsible for modeling the sequential features in both forward and backward temporal directions. Therefore, the proposed model learns the intricate temporal dynamics present in the features obtained from the EEG signals and sensor data. Better classification stability is obtained as a result of a densely connected BiGRU layer. Efficient information flow is guaranteed by the developed AVDBiGRU and it reduces gradient vanishing problems. The hyperparameters of the proposed AVDBiGRU model are optimized by using FFbOA algorithm, while the network weights and biases are trained using the Adam optimizer through standard back propagation for accurate epilepsy classification. The epilepsy analysis efficiency of the implemented model is verified among conventional techniques. The structural design of the developed focal and generalized epilepsy detection model is visualized in Fig. 1.

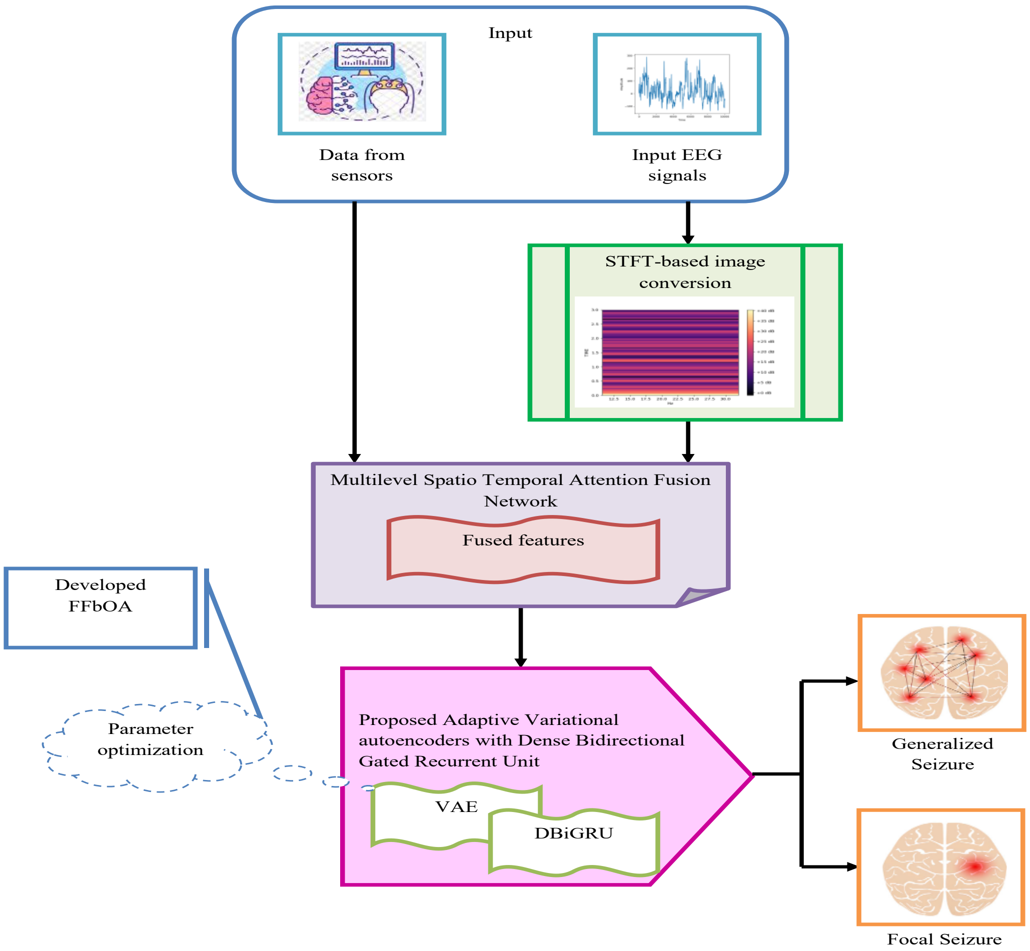


Fig.1. Structural design of the developed focal and generalized epilepsy detection model

3.2. Description of Multi-modal Data

Multimodal data are collected for focal and generalized epilepsy, and their details are provided below:

Sensor Data Dataset Containing HFO Markings for 30 Pediatric Patients with Epilepsy: This database was accessed on 2025-11-08 from https://openneuro.org/datasets/ds003555/versions/1.0.1. EEG signals of 30 adolescents and children are available in this dataset. It contains metadata files in the BIDS standard. These signals are recorded during the sleeping phase of participants. The size of this repository is 15.07 GB. In this repository, samples for different classes including normal, generalized seizure and focal seizure are available. For each class, 25,415 records are available.

The appropriate EEG-based features were accessed on 2025-11-08 from https://www.kaggle.com/datasets/datasetengineer/epilepsy-dataset. Seizure-specific features, wavelet transform features, frequency-domain features, time-domain features and many other types of features are available in this repository. Details like seizure pattern and duration of seizures are given in this dataset. The collected EEG signals and data are specified as S_E and D_S, respectively. The sample signals are shown in Fig. 2.

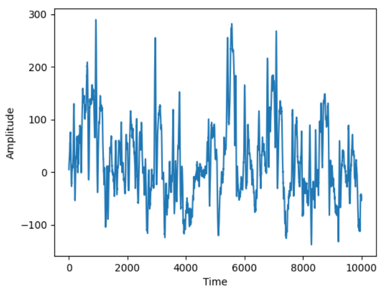


Image 1

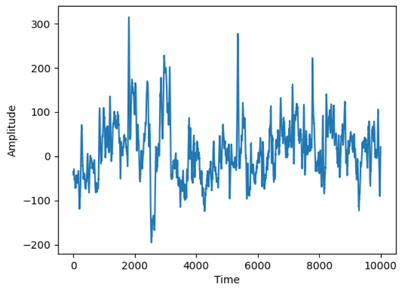


Image 2

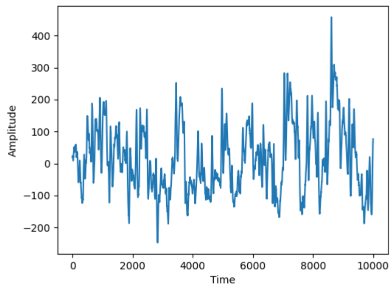


Image 3

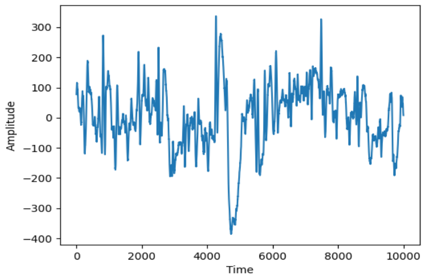


Image 4

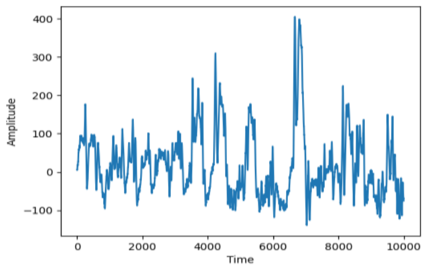


Image 5

Fig.2. Sample EEG signals for focal and generalized epilepsy

3.3. Proposed FFbOA Description

The classification ability of the implemented AVDBiGRU network is enhanced via the FFbOA-based hyperparameter optimization. This algorithm fine-tunes the learning rate as well as hidden neuron counts from the developed network to improve the epilepsy classification accuracy. By optimizing these parameters within a specific range, the proposed model can analyze complex fused features from multimodal data without overfitting. Thus, FFbOA-assisted optimization improves convergence time and increases focal and generalized epilepsy classification accuracy. In the conventional FbOA [26], the ball passing strategies during the football game is mathematically modelled, which helps to enhance its global search. It efficiently escapes local optima by formulating the team tactics during the match. However, the maintenance of the team coordination phase in this algorithm leads to increased computational complexity and it affects the optimization performance when handling large populations. Thus, the performance of FboA is enhanced by modifying the random parameter. In the developed FFbOA, the fitness functions are considered to upgrade the value of random variable d as per Eq. (1).

$$d = \frac{Cntf + Btfn^2}{Cntf + Wstf^2 + Mftn} \tag{1}$$

Here, the current fitness, worst fitness, mean fitness and best fitness are represented as $Cntf$, $Wstf$, $Mftn$ and $Btfn$, concurrently. Upgrading the random number based on the above formulation helps to approach the global optimum quickly. As the random influence is increased in the early stages, the high-dimensional optimization issues are handled precisely by the proposed FFbOA. Due to the consideration of different fitness functions, exploration of various regions in the search space is enhanced. The consideration of random number computation makes the search process more effective across difficult optimization landscapes by improving the convergence speed. The pseudocode of the FFbOA is shown in Algorithm 1.

Algorithm 1: Proposed FFbOA
Input: Initial parameters of VAE and BiGRU
Assign the total iterations B and population F
Compute the fitness of all players
While $(b \le B)$ do
    For $i = 1 : F$
    Perform the exploration phase
    Update the random value as given in Eq. (1)
    Compute the football velocity
    Update the best force

Perform the exploitation phase
Compute the mutation strategy
End for
End while
Return the best solution

Output: Optimized hidden neuron count DR_M^{VAE} and DR_E^{BGRU} and learning rate YR_J^{VAE} and YR_T^{BGRU} from VAE and BiGRU

4. Spectrogram Conversion and Data-Guided Multi-modal Data Fusion Steps for Performance Improvement

4.1. STFT for Spectrogram Conversion

The input EEG signals S_E are converted in to spectrogram images via the STFT approach. The role of STFT [30] is to analyze the changes in frequency of signals over time. Instead of analyzing the entire signal at once, STFT divide the signal into small sections and analyze each part separately. For every segment, a Fourier transform is performed to recognize the frequencies present during the specific time window. STFT helps to examine the interaction between frequency and time and interact. As the signals do not have a consistent frequency, the STFT helps to convert them in an imaging format. The STFT is formulated in Eq. (2).

$$T(\phi,\psi)=\int_{-\infty}^{\infty} t(s).\psi(s-\phi).e^{-i\psi s}\,dt \tag{2}$$

In the above equation, the center of the window function and the time-domain signal is specified as φ and $t(s)$, respectively. The window function placed over φ is given as $\psi(s-\varphi)$. This window function is included by the Fourier transform to analyze the signal segments. In this process, the signal is segmented into overlapping time intervals using a window function. By continuously shifting this window across the entire signal, the progression of frequency over time is analyzed, which results in a time-frequency domain representation. The resultant spectrogram images are symbolized as Z_i^{STFT} and the images are shown in Fig. 3.

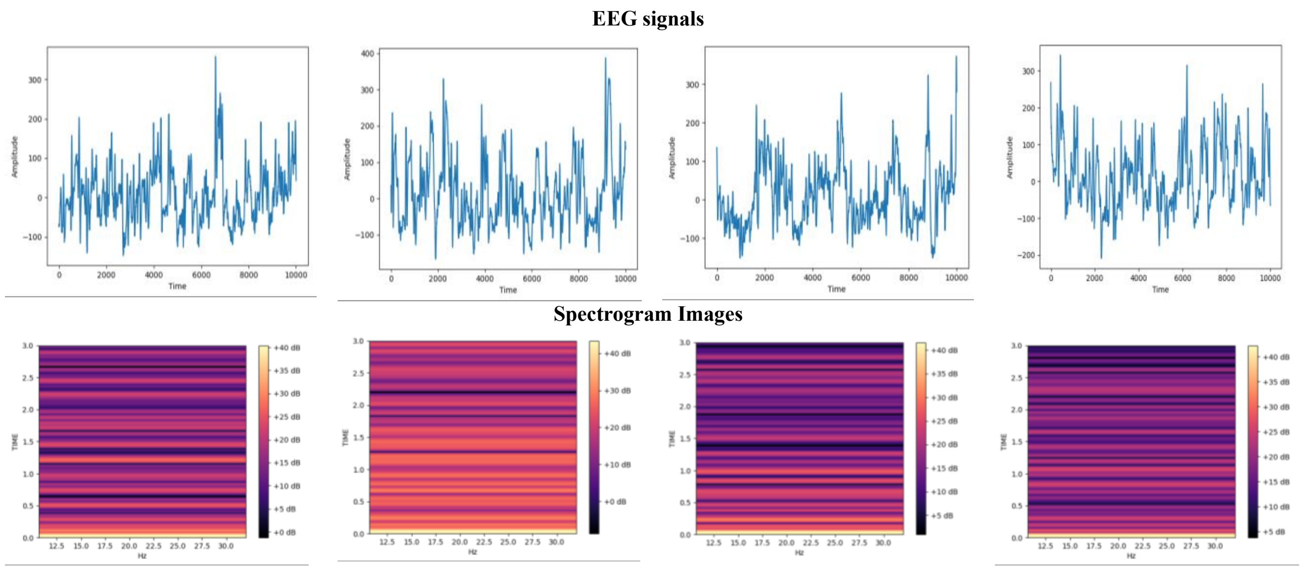


Fig.3. Visualization of resultant spectrogram images

4.2. Spatio Temporal Attention Network

The STA module consists of a self-attention mechanism to learn the long-term spatial and temporal relationships from the spectrogram and sensor data. The STA [31] approach is composed of three main components including the temporal self-attention, spatial self-attention and a feature fusion module. The spatial and temporal attention operates separately to extract respective information by flattening the features. The outputs from these two attention streams are merged through the feature fusion module, which incorporates Multi-Layer Perceptron (MLP) layers, Layer Normalization (LN) and convolution operations.

Initially, the embedded input token $v_m^{'}$ is reshaped. Then the spatial and temporal dimensions generate two separate embedded sequences. The Multi-Head Self-Attention (MSA) processes the flattened sequences to capture the long-range contextual information by managing the spatial and temporal axis. This procedure is mathematically defined in the following Eq. (3)-Eq. (6).

$$A_s^{'} = MSA(LN(A_s)) + A_s \tag{3}$$

$$A_t^{'} = MSA(LN(A_t)) + A_t \tag{4}$$

$$MSA(A) = Concat(h_1, \ldots, h_n)E^l \tag{5}$$

$$h_j = Att(H_j, B_j, P_j) = Soft\max\left(\frac{H_j, B_j^{\mathrm{T}}}{\sqrt{x_b}}\right)P_j \tag{6}$$

In the above notations, the output temporal and spatial features are specified as $A_t^{'}$ and $A_s^{'}$, respectively. The scaling factor is termed as $\sqrt{x_b}$. The linear transformation matrices, attention function and the number of heads is indicated as E^l , Att and h , concurrently. The query, key and value of heads are mentioned as H_j , B_j and P_j , respectively. The attention score between $[0-1]$ is given by the softmax function.

4.3. MSTAFN-based Fusion Process

The input spectrogram images Z_j^{STFT} and data D_S are given to the MSTAFN for retrieving and fusing the features. **Rationale:** Different time points and frequency bands are assigned with high-priority weights by the spatio-temporal attention technique. The EEG data is intrinsically non-stationary and its statistical characteristics change over time. But, the spatio-temporal attention is capable of handling the long-range features from EEG data. Information from the spectrograms and EEG data are easily combined by the MSTAFN model. This feature fusion process eliminates the unnecessary or redundant information. Therefore, MSTAFN is employed for the future fusion process.

Workflow: In the proposed MSTAFN architecture, the fusion is achieved by processing the two input feature sets through separate spatial as well as temporal attention branches and then combining their outputs. The network produces two intermediate feature maps including the spatial feature map and a temporal feature map. The feature and spatial feature maps are converted into a set of query, key, and value representations. The attention mechanism computes the relevance between temporal and spatial positions by multiplying the query and key for generating an attention weight matrix. These weights are then applied to the value representation to produce the temporal as well as spatial attention output. Both these features are concatenated by the feature fusion process of MSTAFN as shown in Eq. (7).

$$FE = MLP\left(LN\left(CF\left(A_s', A_t'\right)\right)\right) + y\hat{A} \tag{7}$$

The output fused feature is signified as FE . A Gaussian Error Linear Unit (GELU) and two fully connected layers are presented in the MLP module. The learnable parameter is symbolized as y . The concatenation process is mentioned by the variable CF . After the fusion, the fused features are obtained based on Eq. (8).

$$FE_v^{ST} = S_E + D_S \tag{8}$$

Insufficient fusion is prevented by preserving the local structural information in the feature fusion module. The multi-level operation captures the hierarchical representation of input, which incorporates the global structures and fine-grained local patterns. The spatio-temporal attention mechanism concentrates on the most informative features associated with epileptic activity. The capability of the MSTAFN-based fusion method to learn the correlations between features from the input images and data is high. Moreover, the suggested technique has the potential to learn the deep contextual interactions across various abstraction levels by using the multi-level attention. The temporal phases of neuronal activity are highlighted by the discriminative ability of the spatio-temporal attention mechanism. Furthermore, the attention-driven fusion process reduces the computational complexity. The fused features are represented as FE_v^{ST} . Schematic view of the MSTAFN-based feature fusion process is displayed in Fig. 4.

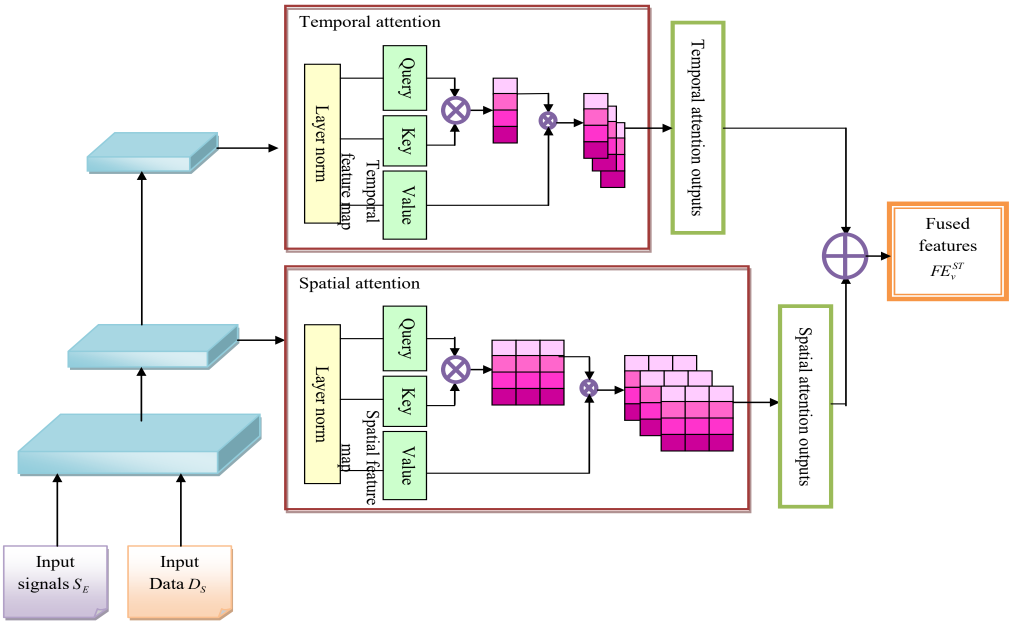


Fig.4. Schematic view of the MSTAFN-based feature fusion process

5. Focal and Generalized Epilepsy Classification using Adaptive Integrated Deep Learning Architecture

5.1. Variational Autoencoders

VAE is employed in this work for reducing the dimensionality of the fused features. An encoder and a decoder are the two main parts of the VAE [32]. The spatial dimension is reduced by the convolutional layers in VAE. One convolution layer has 32 filters and the other layer has 64 filters. The latent space mean is generated via the dense layer and log variance. Eq. (9) provides the log-likelihood of the data.

$$\log d_\theta(v) \geq \mathrm{E}_{k\phi(x|v)}\left[\log d_\theta(v|x)\right] - KL\left[k_\phi(x|v) \| d(x)\right] \tag{9}$$

Here, the Kullback-Leibler divergence and the prior distribution are represented as KL and $d(x)$, respectively. The probability of latent variables and the approximate posterior is termed as $d_\theta(v|x)$ and $k_\phi(x|v)$, concurrently. The fully connected layer in the decoder part reconstructs the input. Therefore, the latent vector is reshaped into a $56 \times 56 \times 32$ tensor. The output is normalized by the transposed convolutional layer. Architectural overview of VAE is depicted in Fig. 5.

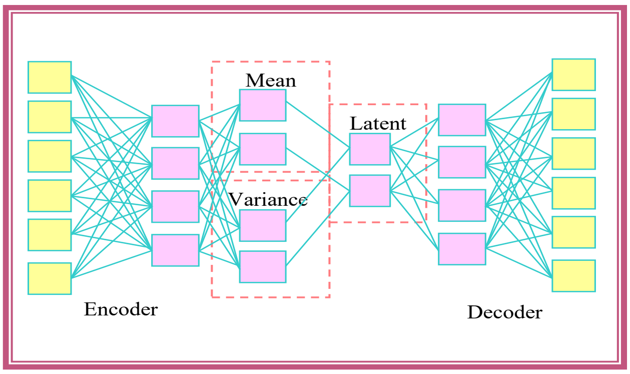


Fig.5. Architectural overview of VAE

5.2. Dense Bidirectional Gated Recurrent Unit

The DBiGRU [33] model performs the epilepsy detection without taking a longer training time. The feature learning efficiency of the BiGRU is enhanced by its densely connected structure. Feature reusing ability is enhanced by the Dense BiGRU. This network prevents gradient issues by passing the information to intermediate layers. In BiGRU, the adjustments needed on the hidden state is managed by the update and reset gates. The information storage is controlled by these gates. The operations inside the GRU component are expressed in the following Eq. (10)-Eq. (12).

$$p_a = \sigma\left(H_p f_a \oplus D_p g_{a-1}\right) \tag{10}$$

$$k_a = \sigma\left(H_k f_a \oplus D_k g_{a-1}\right) \tag{11}$$

$$g_a = \left(\left(1 - p_a\right) \otimes g_{a-1}\right) \oplus \left(p_a \otimes q_a\right) \tag{12}$$

In the above derivations, the prior hidden state, sigmoid, input features, and time are specified as g_{a-1}, σ, f_a and a, respectively. The update and reset gate are mentioned as p_a and k_a, concurrently. In DBiGRU, the contextual data from previous and future time steps are considered for providing the output. Two hidden states are trained in this network. The computations of the BiGRU are given in Eqs. (13)-(15).

$$\overrightarrow{g_a} = GRU\left(f_a, \overrightarrow{g_{a-1}}\right) \tag{13}$$

$$\overleftarrow{g_a} = GRU\left(f_a, \overleftarrow{g_{a-1}}\right) \tag{14}$$

$$g_a = \left[\overrightarrow{g_a}, \overleftarrow{g_a}\right] \tag{15}$$

Here, the forward and backward hidden states are represented as $\overrightarrow{g_a}$ and $\overleftarrow{g_a}$, correspondingly. One of the hidden state processes the inputs forward, while the other processes the input in the opposite direction. The hidden states from both directions are used to calculate the outputs at each time step. The DBiGRU accurately represents the sequential data by utilizing both past and future information. The architectural overview of the DBiGRU is showcased in Fig. 6.

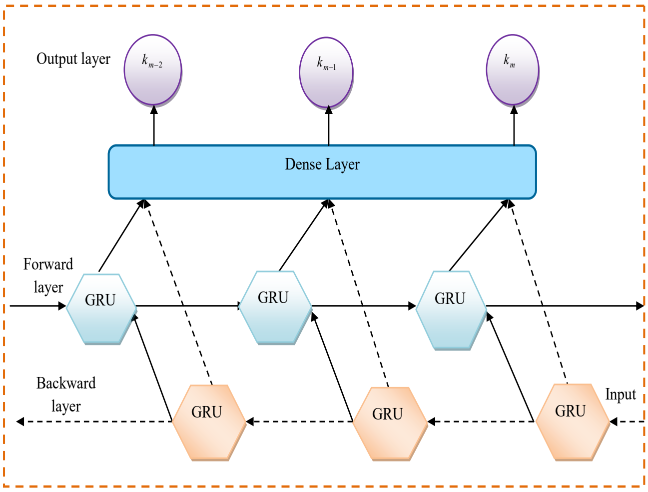


Fig.6. Architectural overview of DBiGRU

5.3. *Epilepsy Classification through AVDBiGRU*

The fused features FE_v^{ST} are processed by the proposed AVDBiGRU technique for classifying focal and generalized epilepsy.

The working principle of the AVDBiGRU is described as follows: The VAE module converts the fused features into a compact latent space. The VAE encoder maps the input to a lower-dimensional latent space by estimating the parameters of the latent distribution. From this latent space, the decoder performs the reconstruction process. VAE identifies the useful characteristics and patterns in the input. The nonlinear relationships between the input and the latent parameters are captured using VAE. The generative ability of VAE reduces the overfitting issues and enhances the learning efficiency for yielding useful latent embeddings for the classification of focal and generalized epilepsy. The VAE generates meaningful representations from the data and denoises the input before further processing. The resultant features from the VAE are classified by the DBiGRU. In the DBiGRU, one set of neurons processes data in the forward direction, while the other processes data in the backward direction. The contextual information from the fused features is learned by the bidirectional structure of the DBiGRU model. The intricate temporal relationships between the epileptic events are precisely handled by the developed AVDBiGRU network. Deeper feature propagation and information reuse between layers are facilitated by the dense connections within the DBiGRU layers. Furthermore, the gating mechanisms in the DBiGRU manage the information flow, which helps to preserve long-term relationships while eliminating unnecessary temporal details. The outputs of the forward and backward DBiGRU units are concatenated to form the temporal features needed for classification. The dense classification layer receives the aggregated temporal features. This layer acts as the decision-making component of the model for determining whether the input pattern corresponds to focal or generalized epilepsy based on the learned latent and temporal features.

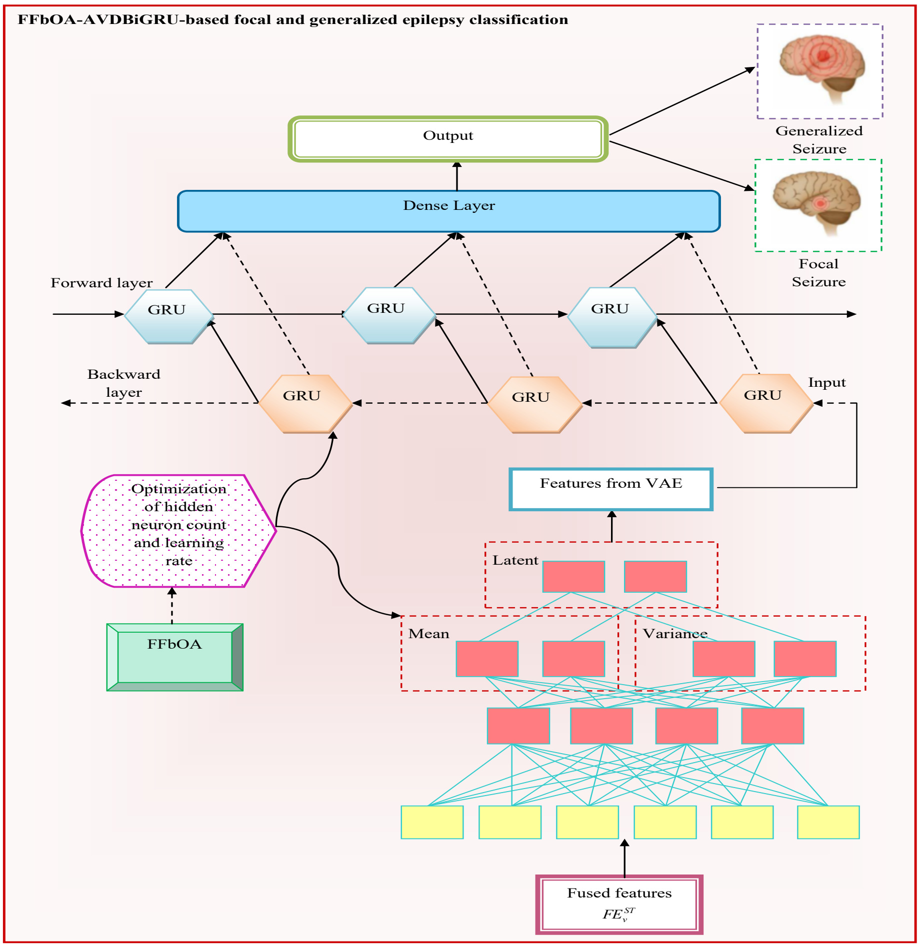


Fig.7. Structural view of the proposed FFbOA-AVDBiGRU-aided focal and generalized epilepsy detection model

The integration of VAE and BiGRU improves the gradient flow and feature learning. The training time of the AVDBiGRU network is minimized by the parameter optimization process by the FFbOA. The objective function Lfn of the implemented FFbOA-AVDBiGRU-aided focal and generalized epilepsy detection model is given in Eq. (16).

$$Lfn = \underset{\left\{GT_J^{XGB}, GT_M^{BiLSTM}, PL_H^{XGB}, PL_A^{BiLSTM}\right\}}{\arg\min} \left(\frac{1}{Accuracy + \Pr ecision} \right) \tag{16}$$

In the above expression, the learning rate optimized from VAE and BiGRU in the range of $[0.01 - 0.99]$ is mentioned as YR_J^{VAE} and YR_T^{BGRU} , respectively. The hidden neuron counts of VAE and BiGRU in the interval of $[5 - 255]$ is represented as DR_M^{VAE} and DR_E^{BGRU} , concurrently. The arithmetical expression for the accuracy and precision is provided in Eq. (17) and Eq. (18).

$$Accuracy = \frac{pp_T + nt_T}{pp_T + nt_T + fp_F + fn_F} \tag{17}$$

$$\Pr ecision = \frac{pp_T}{pp_T + fp_F} \tag{18}$$

In the above equations, the terms fp_F , pp_T , fn_F and nt_T specifies the false positive, true positive, false negative and true negative values. The structural view of the proposed FFbOA-AVDBiGRU-aided focal and generalized epilepsy detection model is exhibited in Fig. 7.

6. Experimental Observations and Discussion

6.1. Experimental Configuration

The proposed multi-modal focal and generalized epilepsy classification model was implemented using Python. During the optimization phase, the number of populations was set to 10, the chromosome length to 4, and the maximum number of iterations to 50. The effectiveness of the proposed FFbOA was compared against Carpet Weaver Optimization (CWO) [27], Dark Forest Algorithm (DFA) [28], Sharpbelly Fish Optimization Algorithm (SBFOA) [29] and FboA [26]. For classification benchmarking, various neural network models such as LSTM [1], DCNN [3], InceptionV3 [4], and BiGRU [33] were evaluated against the proposed model. To ensure an unbiased evaluation, a subject-wise K-Fold cross-validation strategy is used, where each subject's data is kept within a single fold to prevent data leakage and improve the generalizability of the model.

6.2. Assessment Parameters

The computations for the assessment parameters considered to evaluate the performance of the implemented AVDBiGRU-assisted focal and generalized epilepsy detection are provided in Eqs. (19)-(28).

$$\mathrm{F}PR = \frac{fp_F}{fp_F + nt_T} \tag{19}$$

$$\mathrm{F1\text{-}score} = \frac{2pp_T}{2pp_T + fp_F + fn_F} \tag{20}$$

$$\mathrm{F}DR = \frac{fp_F}{fp_F + pp_T} \tag{21}$$

$$CSI = \frac{pp_T}{pp_T + fp_F + fn_F} \tag{22}$$

$$FOR = \frac{fn_F}{fn_F + nt_T} \tag{23}$$

$$FPR = \frac{fp_F}{fp_F + nt_T} \tag{24}$$

$$MCC = \frac{pp_T \times nt_T - fp_F \times fn_F}{\sqrt{(pp_T + fp_F)(pp_T + fn_F)(nt_T + fp_F)(nt_T + fn_F)}} \tag{25}$$

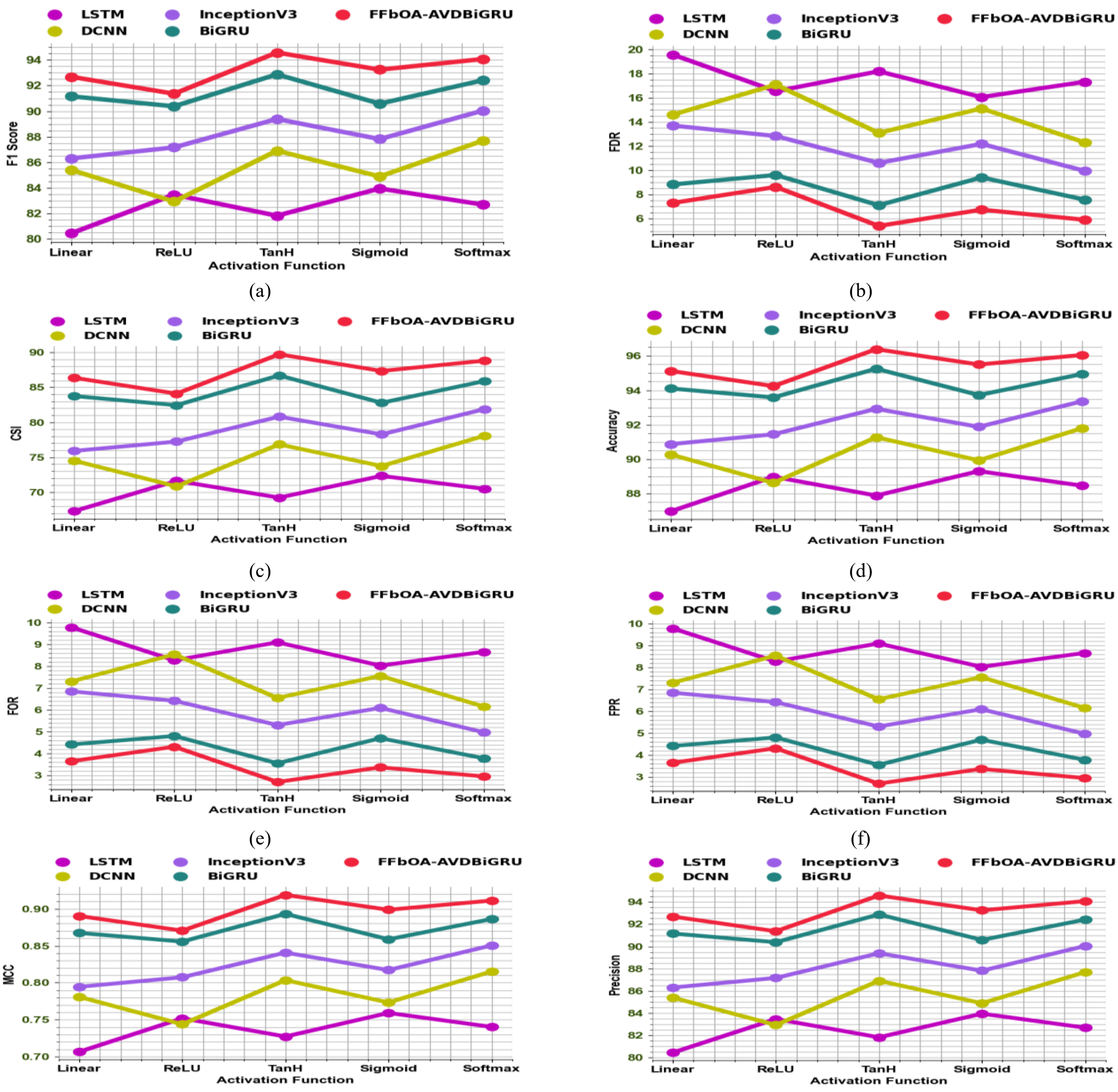
$$FNR = \frac{fn_F}{fn_F + pp_T} \tag{26}$$

$$\mathrm{N}PV = \frac{nt_T}{nt_T + fn_F} \tag{27}$$

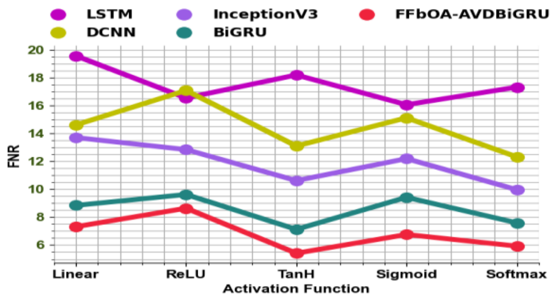
$$\text{Specificity} = \frac{nt_T}{nt_T + fp_F} \tag{28}$$

6.3. Evaluation of Seizure Identification Accuracy

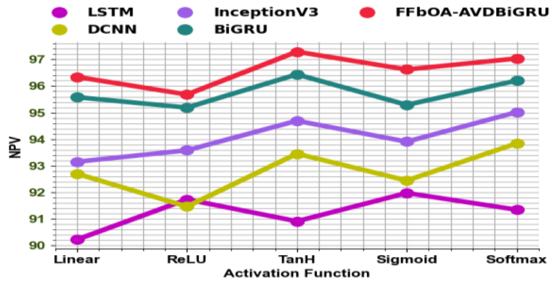
The performance of the proposed model in terms of focal and generalized seizure identification accuracy is compared with conventional approaches, as shown in Figs. 8 and 9. When using the TanH activation function, the F1-score of the developed FFbOA-AVDBiGRU-based epilepsy classification model is 13.75%, 8.46%, 5.29%, and 1.58% higher than those of the LSTM, DCNN, InceptionV3, and BiGRU models, respectively. The MSTAFN efficiently combines multimodal features from sensor-based inputs with STFT-based EEG representations to accurately capture temporal and spatial features, resulting in higher F1-scores and lower FDR values across all activation functions. The precision of the FFbOA-AVDBiGRU-based epilepsy classification is higher than that of the CWO-AVDBiGRU, DFA-AVDBiGRU, SBFOA-AVDBiGRU, and FbOA-AVDBiGRU models. By removing noise and redundancy from multimodal data, the VAE component improves the extraction of enhanced latent representations that help to obtain high accuracy and precision. The data processing capability of the FFbOA-AVDBiGRU in both forward and backward directions enhances the learning of temporal dependencies in epileptic signals. Traditional models like LSTM and BiGRU lack feature-level attention mechanisms and primarily handle sequential learning. Even though techniques like DCNN and InceptionV3 are efficient in spatial learning, they are not well-suited for processing long-term temporal patterns.



(a) (b) (c) (d) (e) (f) (g) (h)

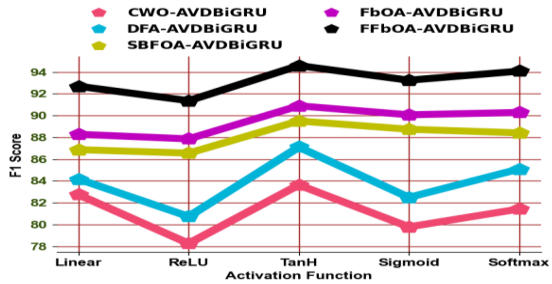


(i)

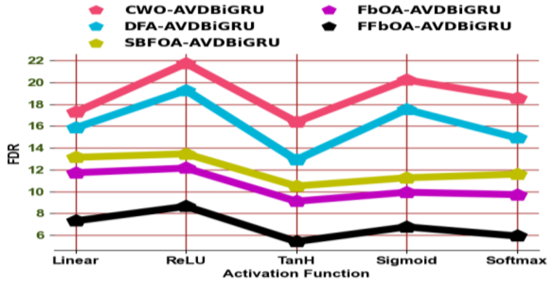


(j)

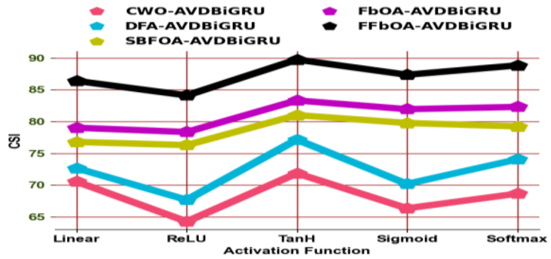
Fig.8. Evaluation of seizure identification performance of the proposed model among different classifiers by means of a) F1-score, b) FDR, c) CSI, d) Accuracy, e) FOR, f) FPR, g) MCC, h) Precision, i) FNR, and j) NPV



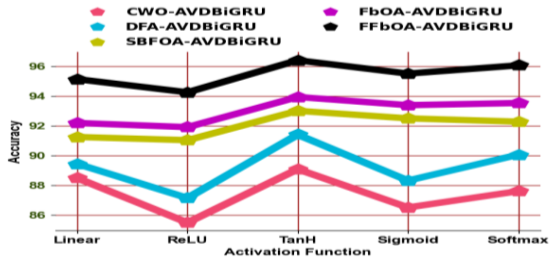
(a)



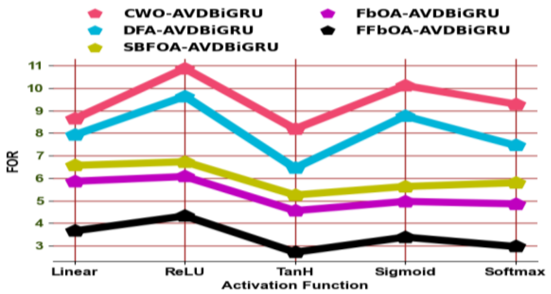
(b)



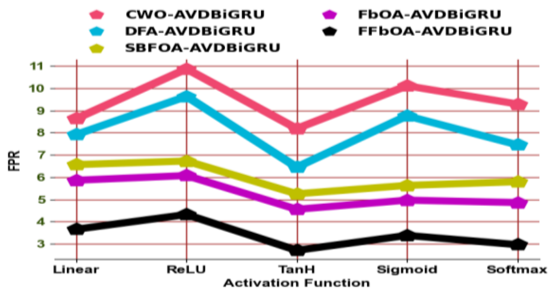
(c)



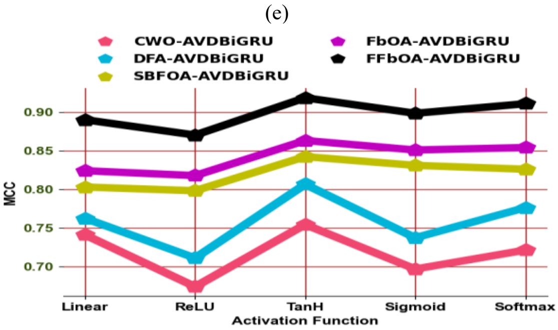
(d)



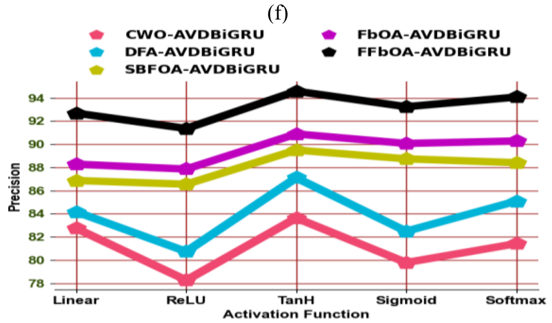
(e)



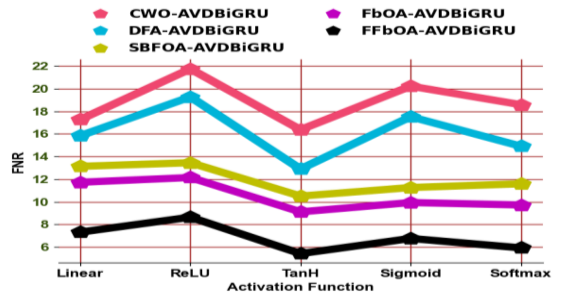
(f)



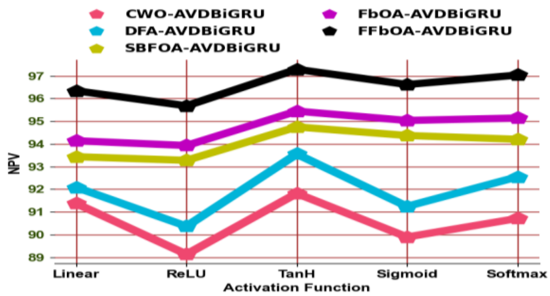
(g)



(h)



(i)



(j)

Fig.9. Evaluation of seizure identification performance of the proposed model among different optimization strategies by means of a) F1-score, b) FDR, c) CSI, d) Accuracy, e) FOR, f) FPR, g) MCC, h) Precision, i) FNR, and j) NPV

6.4. *Quantitative Verification of Model Performance*

The quantitative verification of the implemented model's performance against existing algorithms and classifiers is presented in Tables 2 and 3. These results show the efficiency of the proposed FFbOA-AVDBiGRU network in identifying the focal as well as generalized epilepsy. The accuracy of the designed FFbOA-AVDBiGRU approach is 95.54% with a hidden neuron count of 50, whereas the accuracies of LSTM, DCNN, InceptionV3, and BiGRU are 90.69%, 91.82%, 94.07%, and 95.15%, respectively. The multilayer fusion process preserves high-level temporal features from sensor data and low-level textural features from STFT images with minimal distortion. Additionally, the VAE improves the latent space representation and produces more stable convergence, resulting in higher accuracy, CSI scores, and lower FNR in the detection process. The dense connectivity of the implemented FFbOA-AVDBiGRU network minimizes information loss and improves gradient flow. With 250 epochs, the proposed FFbOA-AVDBiGRU model achieved a specificity score of 96.55%. Meanwhile, algorithms such as CWO-AVDBiGRU, DFA-AVDBiGRU, SBFOA-AVDBiGRU, and FbOA-AVDBiGRU achieved specificities of 90.95%, 93.87%, 93.85%, and 95.73%, respectively. Although LSTM is proficient in capturing sequential dependencies, it struggles with vanishing gradient issues and lacks the ability to recognize important spatial features. InceptionV3 incurs high computational costs and is less efficient in managing temporal changes. Therefore, the performance of the implemented technique in focal and generalized epilepsy detection surpasses that of standard techniques.

Table 2. Quantitative verification of model performance over heuristic algorithms

Hidden neuron count	CWO-AVDBiGRU [27]	DFA-AVDBiGRU [28]	SBFOA-AVDBiGRU [29]	FbOA-AVDBiGRU [26]	FFbOA-AVDBiGRU
Accuracy					
50	89.61257	90.39592	91.5428	92.97386	95.54996
100	86.66371	89.09126	91.63045	93.37707	94.05339
150	87.02448	89.66885	92.19974	93.79781	94.38462
200	88.04126	88.3236	90.59245	91.779	94.67711
250	87.94623	91.8399	91.80391	94.3145	95.41063
Specificity					
50	92.20943	92.79694	93.6571	94.7304	96.66247
100	89.99779	91.81845	93.72284	95.0328	95.54004
150	90.26836	92.25164	94.1498	95.34835	95.78847
200	91.03095	91.2427	92.94434	93.83425	96.00783
250	90.95967	93.87992	93.85294	95.73588	96.55798
FNR					
50	15.58114	14.40612	12.6858	10.53921	6.675063
100	20.00443	16.36311	12.55432	9.934398	8.919921
150	19.46328	15.49672	11.70039	9.303291	8.423063
200	17.93811	17.5146	14.11133	12.3315	7.984333
250	18.08066	12.24015	12.29413	8.528248	6.884048
F1-score					
50	84.41886	85.59388	87.3142	89.46079	93.32494
100	79.99557	83.63689	87.44568	90.0656	91.08008
150	80.53672	84.50328	88.29961	90.69671	91.57694

200	82.06189	82.4854	85.88867	87.6685	92.01567
250	81.91934	87.75985	87.70587	91.47175	93.11595
CSI					
50	73.03861	74.81582	77.48465	80.93127	87.48524
100	66.66052	71.87578	77.69198	81.92668	83.62114
150	67.41546	73.16509	79.0504	82.97711	84.4626
200	69.58047	70.19162	75.26743	78.04445	85.21205
250	69.37575	78.18935	78.1037	84.28382	87.11866

Table 3. Quantitative verification of model performance over techniques

Hidden neuron count	LSTM [1]	DCNN [3]	InceptionV3 [4]	BiGRU [33]	FFbOA-AVDBiGRU
Accuracy					
50	90.69763	91.82329	94.07645	95.15229	95.54996
100	89.6301	90.85079	93.10857	93.62065	94.05339
150	88.3845	90.71332	93.06705	93.30787	94.38462
200	88.92887	89.53507	90.4439	93.51639	94.67711
250	90.58414	92.37136	92.69337	94.26745	95.41063
Specificity					
50	93.02322	93.86747	95.55734	96.36422	96.66247
100	92.22258	93.1381	94.83143	95.21549	95.54004
150	91.28837	93.03499	94.80029	94.9809	95.78847
200	91.69665	92.1513	92.83292	95.13729	96.00783
250	92.93811	94.27852	94.52003	95.70058	96.55798
FNR					
50	13.95355	12.26506	8.885321	7.27157	6.675063
100	15.55485	13.72381	10.33714	9.56902	8.919921
150	17.42326	13.93002	10.39942	10.0382	8.423063
200	16.60669	15.6974	14.33415	9.725413	7.984333
250	14.12379	11.44297	10.95995	8.598832	6.884048
F1-score					
50	86.04645	87.73494	91.11468	92.72843	93.32494
100	84.44515	86.27619	89.66286	90.43098	91.08008
150	82.57674	86.06998	89.60058	89.9618	91.57694
200	83.39331	84.3026	85.66585	90.27459	92.01567
250	85.87621	88.55703	89.04005	91.40117	93.11595
CSI					
50	75.5101	78.14981	83.67949	86.44269	87.48524
100	73.07798	75.86467	81.26262	82.53335	83.62114
150	70.32401	75.54635	81.16037	81.75507	84.4626
200	71.51674	72.86473	74.92586	82.27318	85.21205
250	75.24831	79.464	80.24522	84.16404	87.11866

6.5. Efficiency Analysis of Feature Fusion Process

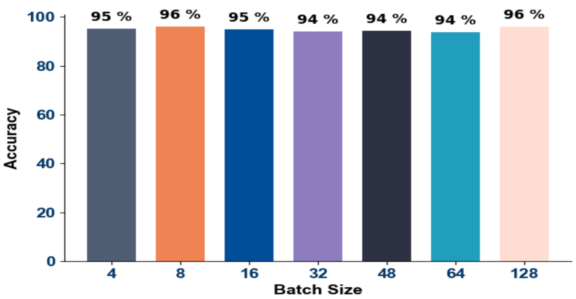
The efficiency of using MSTAFN-based fused features for focal and generalized epilepsy classification is validated, with results presented in Table 4. When using spectrogram images for focal and generalized epilepsy classification, the accuracy is 91.42%. Similarly, when using sensor data for seizure classification, the accuracy is 93.77%. However, using the features obtained by fusing spectrogram and sensor data via the MSTAFN resulted in an accuracy of 95.41%. The MSTAFN module fuses spatial and temporal patterns after processing multi-level feature domains. This fusion strategy effectively integrates the most informative, context-aware, and semantically rich representations, thereby enhancing classification performance. Therefore, the use of MSTAFN-based fused features yielded better precision and F1-scores, as well as lower FNR values.

Table 4. Efficiency analysis of feature fusion process

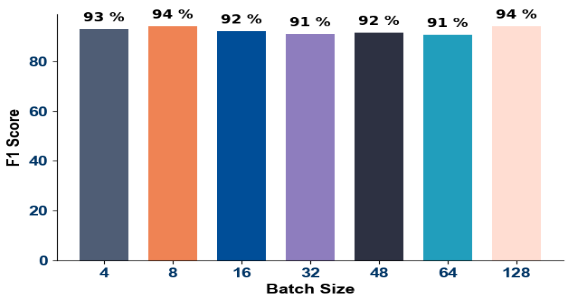
Metrics	Spectrogram images (Feature 1)	Sensor-data(Feature 2)	Fused features using MSTAFN
Accuracy			
Accuracy	91.42447	93.7712	95.41146
Sensitivity	87.13671	90.6568	93.1172
Specificity	93.56836	95.3284	96.5586
Precision	87.13671	90.6568	93.1172
FPR	6.431644	4.671598	3.441402
FNR	12.86329	9.343196	6.882803
NPV	93.56836	95.3284	96.5586
FDR	12.86329	9.343196	6.882803
F1_score	87.13671	90.6568	93.1172
MCC	0.807051	0.859852	0.896758
FOR	6.431644	4.671598	3.441402

6.6. Overall Efficiency of the Presented Model

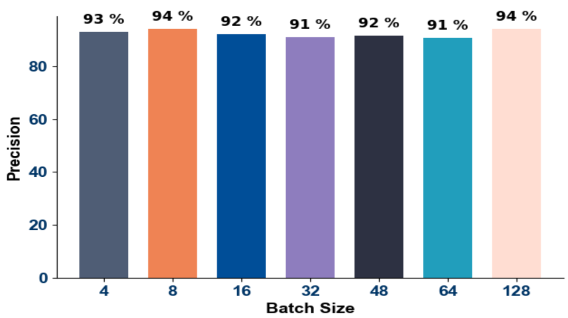
The overall efficiency of the presented FFbOA-AVDBiGRU model for focal and generalized epilepsy classification is analyzed, with results visualized in Fig. 10. The precision analysis graph shows that the proposed model achieved high precision values between 91% and 94% across different batch sizes. The accuracy of 96% is attained by the proposed technique when the batch size is 8 and 128. These results confirm the generalization ability of the developed technique, demonstrating its effectiveness in overcoming gradient instability and overfitting issues. The FNR score of the detection process remains low across different batch sizes, showcasing the efficiency of the FFbOA-based fine-tuning mechanism. The FDR of the developed FFbOA-AVDBiGRU-assisted focal and generalized epilepsy classification is 6% when the batch size is 128 and the FPR is 3%. The proposed model's potential in analyzing the essential feature regions and eliminating unnecessary information resulted in better precision.



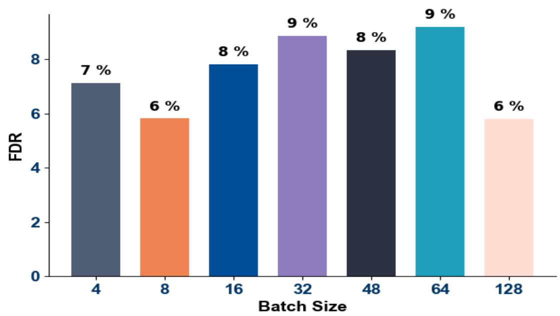
(a)



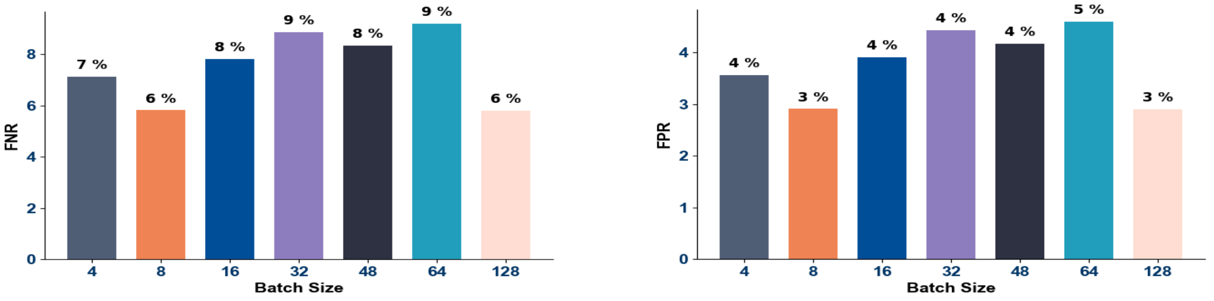
(b)



(c)



(d)

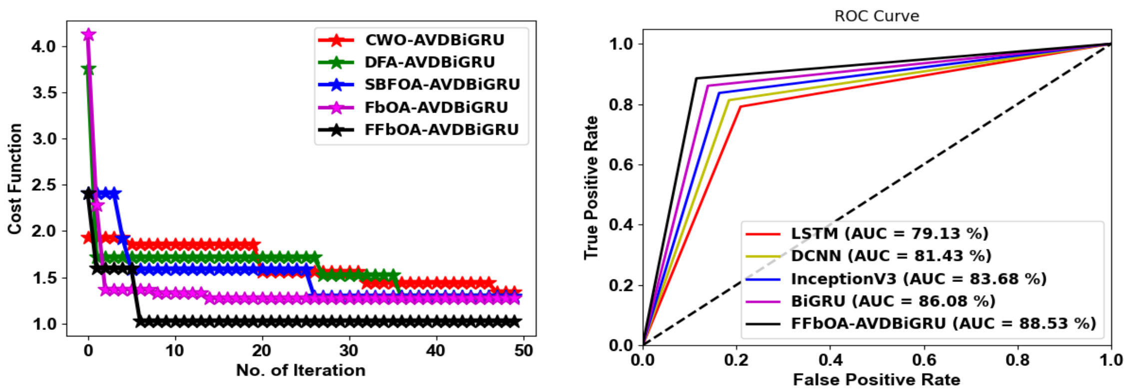


(e) (f)

Fig.10. Overall Efficiency of the Presented Focal and Generalized Epilepsy Classification Model with Respect to a) Accuracy, b) F1-score, c) Precision, d) FDR, e) FNR and f) FPR

6.7. *Convergence and ROC Estimation*

The convergence and ROC estimation of the developed FFbOA-AVDBiGRU model for focal and generalized epilepsy classification is provided in Fig 11. The suggested algorithm's searching ability is confirmed by its reduced cost function value from the very first iteration itself. The convergence outcome indicates that the suggested FFbOA avoids local minima and premature convergence due to the inclusion of fitness-based computation during the optimization process. Therefore, the introduced optimization approach FFbOA enhanced the learning stability of the proposed AVDBiGRU while increasing the accuracy of the detection procedure. The proposed model precisely identified both the focal and generalized epilepsy seizure activities as confirmed by the ROC analysis. The AUC of LSTM in focal and generalized epilepsy classification is 79.13%, DCNN is 81.43%, InceptionV3 is 83.68% and BiGRU is 86.08%, while the proposed model's AUC is 88.53%. The bidirectional recurrent path in the AVDBiGRU network enhances the temporal context awareness by capturing both past and future signal patterns from the fused features, which resulted in the accurate identification of seizure variants.



(a) (b)

Fig.11. Performance Estimation of the Epilepsy Classification Model in terms of a) Convergence and b) ROC

6.8. *Statistical Results*

The statistical results of the proposed FFbOA-AVDBiGRU-aided focal and generalized epilepsy classification are shown in Table 5. FFbOA maintains stable optimization as confirmed by the lowest mean, median, and standard deviation values. The mean and median scores of the implemented approach are 1.114 and 1.03. Traditional algorithms like FbOA and SBFOA show irregular convergence and computational overhead. But, the new concept in the developed FFbOA influences the random factor based on the current, mean, best, and worst fitness values, that enhance the learning stability. Quicker convergence is achieved by the FFbOA-based optimization. Additionally, the FFbOA shows better consistency and control over learning rate as well as hidden neuron count optimization.

Table 5. Statistical results of the presented focal and generalized epilepsy classification model

Statistical measures	CWO-AVDBiGRU [1]	DFA-AVDBiGRU [2]	SBFOA-AVDBiGRU [4]	FbOA-AVDBiGRU [19]	FFbOA-AVDBiGRU
BEST	1.338913	1.291827	1.297593	1.273974	1.022761
WORST	1.928689	3.758084	2.411148	4.123271	2.412587
MEAN	1.636779	1.60502	1.521292	1.369498	1.114488
MEDIAN	1.563311	1.716155	1.588178	1.273974	1.030536
STD	0.203608	0.357632	0.304176	0.418379	0.251429

6.9. Interpretability and Explainabilty of Proposed Model

The attention mechanism in the proposed MSTAFN framework provides the interpretability by highlighting discriminative temporal–spectral EEG regions having contribution in seizure classification, while saliency-based analysis can further identify the most influential input features that are responsible for the prediction.

6.10. Ablation Study of Proposed Model

An ablation study for the proposed model is displayed in Table 6. The final proposed model with used optimization function achieved the highest accuracy of 95.55% and specificity of 96.66% which is improved as compared to the other modules used in proposed model and discussed in ablation study. Although we observed a slight variation in F1-Score and FNR values but the improved specificity and accuracy of proposed model indicates a better class discrimination that is considered decisive in the consistent clinical diagnosis application. Overall, the results validate the effectiveness of each module included in the proposed framework.

Table 6. Ablation study of proposed model

Model	Accuracy (%)	Specificity (%)	FNR (%)	F1-Score (%)
STFT+BiGRU	92.75	92.30	8.12	91.76
STFT+VAR+BiGRU	93.87	93.26	7.17	92.87
STFT+MSTAFN+BiGRU	94.78	94.23	6.26	93.85
STFT+MSTAFN+VAE+BiGRU	95.36	95.15	5.25	94.73
Proposed Model	**95.55**	**96.66**	**6.67**	**93.32**

7. Conclusions

An automatic and effective identification of focal and generalized epilepsy using deep learning was suggested in this work. Initially, the EEG signals and the appropriate sensor data were collected. The input EEG signals were converted into STFT spectrogram images. An MSTAFN was used for retrieving the spatial and temporal features from the EEG as well as sensor data. The multilevel STA model was beneficial in reducing noise. An AVDBiGRU network was proposed for the focal and generalized epilepsy classification. This model processes the fused features obtained from the MSTAFN. The parameters of AVDBiGRU are fine-tuned by the FFbOA for accurate classification of epilepsy. The epilepsy analysis efficiency of the implemented model was verified among conventional techniques. The accuracy of the implemented AVDBiGRU model for focal and generalized epilepsy detection is 95.54%. Thus, the developed technique has high discriminative capacity and it achieved precise classification results. In the future, Empirical Mode Decomposition (EMD) will be used for feature generation instead of only using the STFT mechanism. Noise handling will also be performed via filtering and normalization techniques.

All the Declarations and Statements

Author Contributions Statement

Mr. Maneesh Kumar: Conceptualization, methodology, investigation, data curation, software implementation, formal analysis, visualization, writing – original draft, and correspondence. He proposed the research idea, designed the overall study framework, carried out the implementation and analysis, prepared the manuscript draft, and served as the corresponding author.

Dr. Rakesh Kumar: Supervision, conceptual guidance, validation, review and editing, and project administration. He supervised the research work, provided intellectual direction, reviewed the methodology and results, and contributed to revising and improving the manuscript.

Dr. Santosh Kumar: Supervision, methodology support, validation, review and editing, and technical guidance. He assisted in shaping the research design, supported the evaluation and interpretation of the results, and contributed to the critical review and refinement of the manuscript.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Funding Declaration
This research did not receive any specific grant from funding agencies in the public, commercial, or not for profit sectors.

Data Availability Statement
The authors confirm that the data supporting the finding of this study are available within the article. Raw data that support the findings of this study are available from the corresponding author, upon reasonable request.

Ethical Declarations
This article does not contain any studies with human participants or animals performed by any of the author; hence ethical approval was not required.

Acknowledgments
The authors sincerely thank all those who provided valuable guidance, support, and constructive suggestions during the course of this research. Their insights helped improve the quality and clarity of this work.

Declaration of Generative AI in Scholarly Writing
In the preparation of this manuscript, generative AI and AI-assisted technologies were used solely for language enhancement, grammar correction, and improvement of readability. Their use was limited to the writing process and did not extend to the analysis of data, interpretation of findings, or drawing of scientific conclusions. All AI-assisted output was carefully reviewed, revised, and approved by the authors. The authors remain fully responsible for the accuracy, integrity, and originality of the content presented in this work.

Abbreviations
The following abbreviations are used in this manuscript:

EEG - electroencephalogram
STFT - Short-Time Fourier Transform
MSTAFN - Multilevel Spatio Temporal Attention Fusion Network
AVDBiGRU - Adaptive Variational autoencoders with Dense Bidirectional Gated Recurrent Unit
FFbOA - Fitness-based Football Optimization Algorithm
HFO - High-Frequency Oscillation
ASD - Anti-Seizure Drugs
SOZ - Seizure Onset Zone
CWT - Continuous Wavelet Transform
CNN - Convolutional Neural Networks
LSTM - Long Short-Term Memory Networks
ViT - Vision Transformer
STA - Spatio Temporal Attention Network
MLP - Multi-Layer Perceptron
LN - Layer Normalization
MSA - Multi-Head Self-Attention
GELU - Gaussian Error Linear Unit
VAE - Variational Autoencoders
CWO - Carpet Weaver Optimization
DFA - Dark Forest Algorithm
SBFOA - Sharpbelly Fish Optimization Algorithm

References

[1] Najafi, T., Jaafar, R., Remli, R., and W. A. Wan Zaidi, "A classification model of EEG signals based on RNN-LSTM for diagnosing focal and generalized epilepsy," Sensors, vol. 22, no. 7269, 2022.
[2] Fan, D., Qi, L., Hou, S., Wang, Q., and G. Baier, "The seizure classification of focal epilepsy based on the network motif analysis," Brain Research Bulletin, vol. 207, no. 110879, 2024.
[3] Gill, T. S., Zaidi, S. S. H., and M. A. Shirazi, "Attention-based deep convolutional neural network for classification of generalized and focal epileptic seizures," Epilepsy & Behavior, vol. 155, no. 109732, 2024.
[4] Narin, A., "Detection of focal and non-focal epileptic seizure using continuous wavelet transform-based scalogram images and pre-trained deep neural networks," Irbm, vol. 43, pp. 22-31, 2022.
[5] M. O. Mokhiamar, A. Mahmoud, M. I. Eldagla, Y. Aribi, and A. M. Anter, "Blockchain-integrated multi-modal LSTM-CNN fusion for high-precision epileptic seizure detection from EEG signals," Knowledge-Based Systems, vol. 323, p. 113703, Jul. 2025.
[6] T. Hong, D. Li, Y. Chang, X. Wang, Z. Cai, R. Wang, X. Zhang, X. Liu, C. Yang, S. Yu, S. Liu, and D. Ming, "MCAN: Cross-domain self-supervised attention network based on multiscale EEG feature learning for epileptic seizure detection," Knowledge-Based Systems, vol. 327, p. 114199, Oct. 2025.
[7] Q. Li, W. Cao, and A. Zhang, "Multi-stream feature fusion of vision transformer and CNN for precise epileptic seizure detection from EEG signals," Journal of Translational Medicine, vol. 23, no. 871, 2025.
[8] X. Yi, "Energy-efficient EEG-based epileptic seizure detection: A deep learning-based spiking neural network approach," Procedia Computer Science, vol. 266, pp. 908-915, 2025.
[9] M. Hosseinzadeh, P. Khoshvaght, S. Sadeghi, P. Asghari, A. N. Varzeghani, and M. Mohammadi, "A model for epileptic seizure diagnosis using the combination of ensemble learning and deep learning," IEEE Access, vol. 12, pp. 137132-137143, 2024.
[10] C. Lucasius, V. Grigorovsky, H. Nariai, A. S. Galanopoulou, J. Gursky, and S. L. Moshé, "Biomimetic deep learning networks with applications to epileptic spasms and seizure prediction," IEEE Transactions on Biomedical Engineering, vol. 71, no. 3, pp. 1056-1067, Mar. 2024.
[11] M. J. Rivera, J. Sanchis, O. Corcho, M. A. Teruel, and J. Trujillo, "Evaluating CNN methods for epileptic seizure type classification using EEG data," IEEE Access, vol. 12, pp. 75483-75495, 2024.
[12] V. V. Narayana and P. Kodali, "Hybrid LPF-LSTM model for enhanced epileptic seizure detection in EEG signals," IEEE Sensors Letters, vol. 9, no. 5, pp. 1-4, May 2025.
[13] K. K. Patro, A. J. Prakash, J. P. Sahoo, S. Routray, A. Baihan, and N. A. Samee, "SMARTSeiz: Deep learning with attention mechanism for accurate seizure recognition in IoT healthcare devices," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 7, pp. 3810-3818, Jul. 2024.
[14] Y. Du, Y. Ren, N. Wong, and E. C. H. Ngai, "Hyperdimensional computing with multiscale local binary patterns for scalp EEG-based epileptic seizure detection," IEEE Internet of Things Journal, vol. 11, no. 15, pp. 26046-26061, Aug. 1, 2024.
[15] T. Premavathi, M. Shukla, and R. Jadeja, "Predicting patient-specific epileptic seizures from scalp EEG signals using KNN model with transfer learning," IEEE Access, vol. 12, pp. 196728-196739, 2024.
[16] A. D. Brabandere, C. Chatzichristos, W. Van Paesschen, M. De Vos, and J. Davis, "Detecting epileptic seizures using hand-crafted and automatically constructed EEG features," IEEE Transactions on Biomedical Engineering, vol. 71, no. 1, pp. 318-325, Jan. 2024.
[17] K. Kamakshi and A. Rengaraj, "Early detection of stress and anxiety based seizures in position data augmented EEG signal using hybrid deep learning algorithms," IEEE Access, vol. 12, pp. 35351-35365, 2024.
[18] Y. Ma, J. Wang, J. Wang, W. Kang, and X. Yang, "SyncLearnNet: Generalized epileptic seizure detection network based on brain signals," IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 9, pp. 6565-6575, Sep. 2025.
[19] B. C. Sahoo, S. Mishra, S. K. Satapathy, and S. Kumar, "Robust seizure prediction model using EEG signal for temporal lobe epilepsy leveraging deep learning and continual learning," IEEE Access, vol. 13, pp. 138083-138094, 2025.
[20] S. Baghersalimi, T. Teijeiro, A. Aminifar, and D. Atienza, "Decentralized federated learning for epileptic seizures detection in low-power wearable systems," IEEE Transactions on Mobile Computing, vol. 23, no. 5, pp. 6392-6407, May 2024.
[21] M. G. Murariu, F. R. Dorobanțu, and D. Tărniceriu, "Enhanced classification of focal and generalized epilepsy using EEMD and CEEMDAN methods," Traitement du Signal, vol. 41, no. 1315, 2024.
[22] G. Amrani, A. Adadi, and M. Berrada, "An attention mechanism-based interpretable model for epileptic seizure detection and localization with self-supervised pre-training," IEEE Access, vol. 13, pp. 60213-60232, 2025.
[23] J. Xu, S. Yuan, J. Shang, J. Wang, K. Yan, and Y. Yang, "Spatiotemporal network based on GCN and BiGRU for seizure detection," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 4, pp. 2037-2046, Apr. 2024.
[24] P. Zhu, W. Zhou, C. Cao, G. Liu, Z. Liu, and W. Shang, "A novel SE-TCN-BiGRU hybrid network for automatic seizure detection," IEEE Access, vol. 12, pp. 127328-127340, 2024.
[25] B. Kim, Q. Huang, B. Taylor, Q. Zheng, J. Ku, and Y. Chen, "MulPi: A multi-class and patient-independent epileptic seizure classifier with co-designed input-stationary computing-in-SRAM," IEEE Transactions on Biomedical Circuits and Systems, vol. 19, no. 4, pp. 756-766, Aug. 2025.
[26] M. El-Sayed El-Kenawy, F. H. Rizk, A. M. Zaki, M. E. Mohamed, A. Ibrahim, A. A. Abdelhamid, N. Khodadadi, E. M. Almetwally, and M. M. Eid, "Football optimization algorithm (FbOA): A novel metaheuristic inspired by team strategy dynamics," Journal of Artificial Intelligence and Metaheuristics (JAIM), vol. 8, no. 1, pp. 21-38, 2024.
[27] S. Alomari, K. Kaabneh, I. AbuFalahah, S. Gochhait, I. Leonova, Z. Montazeri, and M. Dehghani, "Carpet weaver optimization: A novel simple and effective human-inspired metaheuristic algorithm," International Journal of Intelligent Engineering & Systems, vol. 17, no. 4, 2024.
[28] D. Li, S. Du, Y. Zhang, and M. Zhao, "Dark forest algorithm: A novel metaheuristic algorithm for global optimization problems," Computers, Materials & Continua, vol. 75, no. 2, 2023.

[29] J. Liu, R. Wang, Y. Deng, X. Huang, and Z. Li, "Sharpbelly fish optimization algorithm: A bio-inspired metaheuristic for complex engineering," Biomimetics, vol. 10, no. 7, p. 445, 2025.
[30] D. Y. Yun and H. S. Park, "Noise-robust structural response estimation method using short-time Fourier transform and long short-term memory," Computer-Aided Civil and Infrastructure Engineering, vol. 40, no. 7, pp. 859-878, 2025.
[31] Z. Han, C. Zhang, L. Gao, Z. Zeng, B. Zhang, and P. M. Atkinson, "Spatio-temporal multi-level attention crop mapping method using time-series SAR imagery," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 206, pp. 293-310, 2023.
[32] F. O. Isinkaye, M. O. Olusanya, and A. A. Akinyelu, "A multi-class hybrid variational autoencoder and vision transformer model for enhanced plant disease identification," Intelligent Systems with Applications, vol. 26, p. 200490, 2025.
[33] Mekruksavanich, Sakorn, and Anuchit Jitpattanakul, "Effective detection of epileptic seizures through EEG signals using deep learning approaches," Machine Learning and Knowledge Extraction, vol. 5, no. 4, pp.1937-1952, 2023.

Authors' Profiles

Maneesh Kumar is a Ph.D. scholar with the Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur, India. His research interests include machine learning, deep learning, artificial intelligence, and data-driven intelligent systems. His research work emphasizes the design and development of advanced computational models for real-world applications. He is actively engaged in exploring innovative methodologies in intelligent computing and interdisciplinary applications of modern AI techniques.

Prof. (Dr.) Rakesh Kumar is Professor and Head of the Department of Computer Science and Engineering at Madan Mohan Malaviya University of Technology, Gorakhpur, India. He received his B.E. in Computer Engineering from Madan Mohan Malaviya Engineering College, Gorakhpur, in 1990, his M.E. in Computer Engineering from SGS Institute of Technology and Science, Indore, in 1994, and his Ph.D. from IIT Roorkee in 2011. Prior to his current position, he served at HBTI Kanpur and BIET Jhansi. He has completed one major UGC-funded research project and two AICTE-sponsored MODROBS projects, and received the Best Research Paper Award at ICIP 2007, Bangalore, India. He was listed among Stanford University's Top 2% Most Influential Scientists in 2023. He has supervised 15 Ph.D. scholars and is currently guiding M.Tech. and Ph.D. students. He has published over 125 research papers in reputed journals and conferences. His research interests include artificial intelligence, IoT, wireless sensor networks, network security, machine learning and data analytics, cloud computing, and image processing.

Dr. Santosh Kumar is a data science researcher specializing in artificial intelligence, machine learning, and data-driven analytics. He is currently associated with the NHS (National Health Service, UK). He holds a Ph.D. in Machine Learning and Artificial Intelligence and an M.Tech in Computer Science and Engineering. His research focuses on applied machine learning, predictive modelling, deep learning, and intelligent decision-support systems, with particular interest in developing computational models and advanced analytical methods for real-world interdisciplinary applications. He also serves as a reviewer for international peer-reviewed journals and conferences in AI, machine learning, and data science.

How to cite this paper

Maneesh Kumar, Rakesh Kumar, Santosh Kumar, "Implementing a Novel Framework for Focal and Generalized Epilepsy Classification using Adaptive VAE with Dense Bi-GRU through Multimodal Data-guided Feature Fusion", International Journal of Intelligent Systems and Applications(IJISA), Vol.18, No.4, pp.183-204, 2026. DOI:10.5815/ijisa.2026.04.10