Explainable classification of seizures and other patterns of harmful brain activity in critically ill patients

Manikandan Arunachalam[1], Sanjay Thiyagarajan[2], Nagandla Chirudeep[3]
[1]Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
[2]Artificial Intelligence and Machine Learning, Birla Institute of Technology and Science (BITS) Pilani, Pilani, India
[3]Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

Article Info

Article history:

Received Mar 1, 2025
Revised Apr 15, 2026
Accepted Jul 14, 2026

Keywords:

Classification
Deep neural networks
Explainable artificial intelligence
Patient monitoring
Seizures

ABSTRACT

Accurate detection and classification of seizures from electroencephalography (EEG) signals play a vital role in enabling timely medical interventions and treatments for neurological disorders. Currently, EEG recordings are analyzed exclusively by trained human experts. Although essential, this manual process is time-intensive, costly, and prone to fatigue-related errors, in addition to inconsistencies between reviewers. In this work, we propose a deep neural network (DNN) model equipped with interpretable layers designed to classify seizures and other abnormal brain patterns, such as periodic discharges, rhythmic delta activity, and miscellaneous pathological events. The proposed DNN architecture incorporates explainable components that allow clinicians to trace and understand the model's reasoning process, fostering confidence and aiding clinical decision-making. This combination of deep learning (DL) with interpretability mechanisms is novel and overcomes several limitations of traditional approaches. The method is validated using a publicly available EEG dataset, achieving state-of-the-art accuracy while maintaining interpretability that supports expert review. This study contributes to the growing field of EEG-based seizure detection by bridging the gap between high-performance DNN-based models and the need for clinical transparency. Unlike existing methods, our approach employs an ensemble of three specialized DNN, each capturing complementary feature representations. This ensemble framework improves both interpretability and robustness of the classification results.

Corresponding Author:

Manikandan Arunachalam
Department of Electronics and Communication Engineering, Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Coimbatore, India
Email: a_manikandan@cb.amrita.edu

1. INTRODUCTION

The precise classification of seizures and other irregular brain movement patterns in critically ill patients is vital for timely diagnosis, suitable treatment, and better neurological outcomes. Patients in intensive care units (ICUs) are mainly vulnerable to various neurological disorders, including seizures, periodic discharges, and rhythmic patterns, which, if left unrecognized or misconstrued, can lead to long-term cognitive damage or even impermanence [1], [2]. Seizures are serious threats to the neurological well-being and are associated with long term cognitive deficits if missed or mistaken [3]. Numerous studies over the past decade have absorbed on the implementation of continuous electroencephalography (cEEG) observing and automated seizure detection systems in ICU settings. A primary machine learning (ML)-based approach was established to perceive seizure events using raw EEG signals. Similarly, use of quantitative EEG features in real-time classification of non-convulsive seizures (NCS) discovered [4]. While these aids laid the groundwork for automated seizure detection, several boundaries persist—such as poor generalizability across diverse patient populations, problems in distinguishing pathological patterns from artifacts, and imperfect interpretability of black-box algorithms [5].

In spite of advances in neuroimaging, cEEG, and algorithmic analysis, present keys fight with robust seizure classification under ICU-specific challenges, including artifact contamination, multi-etiologic neurological conditions, and imperfect data interpretability for clinical use [6]. These questions hamper consistent decision-making in high-stakes situations. Moreover, numerous existing models flop to address the clinical necessity for transparent and explainable artificial intelligence (XAI), which is serious for physician faith and integration into tedious workflows [7]. Seizures in the critically ill frequently existing as non-convulsive or subclinical events, making clinical observation alone inadequate for diagnosis. Research have shown that up to 20% of comatose ICU patients experience NCS or NCS epilepticus (NCSE) [8]. These actions, if unnoticed, can affect in secondary brain damage, protracted hospital stays, and worse neurological conclusions. EEG monitoring has thus become a foundation of neurocritical care, enabling real-time investigation of brain action [9], [10]. However, cEEG data is high-dimensional and complex, necessitating skilled neurologists for interpretation—a source often imperfect in real-time ICU settings. As such, there has been a rising mandate for automated systems proficient of precise and timely classification of seizures and other EEG abnormalities such as intervallic releases and rhythmic delta action [11].

Early exertions in automated seizure detection were rule-based and intensive on handcrafted structures such as spike frequency, amplitude, and rhythmicity. While interpretable, these systems required sturdiness and wriggled with generalization across various patient populations. Additionally, their performance worsened in the existence of noise and artifacts—a common incidence in ICU situations due to patient movement, electrical interference, and poor electrode contact [12]. With the growth of ML, researchers instigated applying classifiers such as support vector machines (SVM), random forests, and k-nearest neighbors (k-NN) to EEG data [13]-[15]. These approaches classically depend on time-domain, frequency-domain, and time-frequency features obtained from EEG signals. For example, SVM-based system established to achieve hopeful results in detecting seizures in pediatric patients [16]. Despite developments in accuracy, these models often essential widespread feature engineering and still required interpretability. Moreover, they were not well-adapted to complex, multi-channel, and high frequency cEEG data common in adult ICU sceneries [17].

The application of deep learning (DL) patented a significant dive in seizure classification research. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been used to study structures directly from raw EEG signals, removing the necessity for manual feature extraction [18]. For instance, a CNN-LSTM hybrid model developed to perceive seizures from multi-channel EEG recordings in the TUH EEG Seizure Corpus. DL models have established superior performance in seizure detection responsibilities, predominantly in terms of sensitivity and specificity. However, these enhancements come at the cost of explainability. Most DL models are perceived as "black boxes," making it hard for clinicians to understand the foundation of their decisions—a significant barrier to clinical adoption. The absence of transparency in DL systems has been extensively documented as a critical restriction, particularly in high-stakes medical applications [19], [20].

XAI intentions to report this by providing human-understandable understandings into model predictions [21]. In the context of seizure detection, XAI can support clinicians verify that a model's decision is based on valid EEG patterns rather than objects or spurious connections [22]. Several XAI techniques have been deployed to EEG data. Saliency maps, layer-wise relevance propagation (LRP), and SHapley Additive exPlanations (SHAP) are among the most commonly used [23]. These methods allow conception of which EEG channels, frequencies, or time segments donated most to the model's decision. An interpretable CNN model projected for seizure detection that uses gradient-weighted class activation mapping (Grad-CAM) to focus important EEG regions [24]. Also, a combined attention mechanisms with CNN architectures was anticipated to improve both performance and interpretability. Another method is to enterprise inherently interpretable models, such as decision trees or attention-based networks, even if these often loss performance. Hybrid models that poise accuracy and explainability are gaining traction [25]. A two-stage model developed where the first stage senses potential seizure segments using a high-sensitivity deep model, and the second stage offers interpretable decisions using rule-based logic. In spite of these advancements, several challenges remain. First, most models are skilled and tested on benchmark datasets like CHB-MIT or TUH EEG, which may not take a broad view to ICU settings with noisier data [26]. Second, inter-patient inconsistency in seizure morphology limits the generalizability of trained models [27]. Third, the trade-off between model accuracy and explainability is not yet well-determined, with many explainable models disappointing in real-world testing. Moreover, evaluation system of measurement in many studies emphasis only on detection accuracy and neglect clinical significance, such as timeliness of detection and false alarm rates [28]. These factors are serious in ICU environments, where clinical workflows be contingent on timely and actionable alerts.

To address these gaps, future research should focus on, emerging standardized, annotated ICU-specific EEG datasets, making hybrid models that syndicate DL with interpretable layers, mixing clinical knowledge and contextual data into model inputs, designing human-in-the-loop systems where AI augments, rather than substitutes, clinician judgment and creating assessment protocols that reflect clinical usability, such as alarm fatigue and decision impact. Explainability should also be tailor-made to end-user needs. For example, while a neurologist may favour thorough signal-level clarifications, ICU staff may profit more from modest, high-level visual summaries or risk scores. In summary, the literature discloses a clear trajectory from rule-based systems to black-box DL models, and more recently, to explainable AI in seizure classification. While performance has progressively enhanced, challenges related to real-time applicability, robustness, and transparency continue. This study figures upon prior work by proposing an explainable classification framework for seizures and abnormal brain activity patterns in critically ill patients. The model integrates state-of-the-art DL for accuracy, along with feature importance analysis and LRP for interpretability. By focusing specifically on ICU EEG data and including a clinician-oriented explanation interface, this work goals to bridge the gap between algorithmic performance and clinical trust. These techniques permit clinicians to well understand the basis behind model decisions, fostering poise in its clinical deployment. Accurate identification and classification of harmful brain activity, mainly seizures, in critically ill patients has developed as a vital factor of neurocritical care. The growing prevalence of neurological complications in ICUs, shared with the complexity of EEG patterns and the need for sensible clinical decision-making, has driven momentous advancements in automated seizure detection and classification skills. While substantial development has been made, key openings continue in terms of generalizability, robustness to artifacts, and most judgmentally, explainability of the models used in clinical decision support systems. This literature review fuses the body of research focused on EEG-based seizure classification in the critically ill, evaluates the evolution and limitations of ML and DL methods applied in this domain, and identifies the critical need for interpretable AI solutions.

2. PROPOSED MODEL

The proposed preprocessing pipeline (Figure 1) is intended to convert raw EEG recordings into a organized, high-quality format suitable for DL-based seizure and abnormal brain activity classification. It integrates multiple sequential stages aimed at refining data consistency, reducing noise, and augmenting training samples to enhance model generalization. The process instigates with file reading and raw data decoding, where EEG signals attained via scalp electrodes are transformed from their native acquisition format into a structured representation.

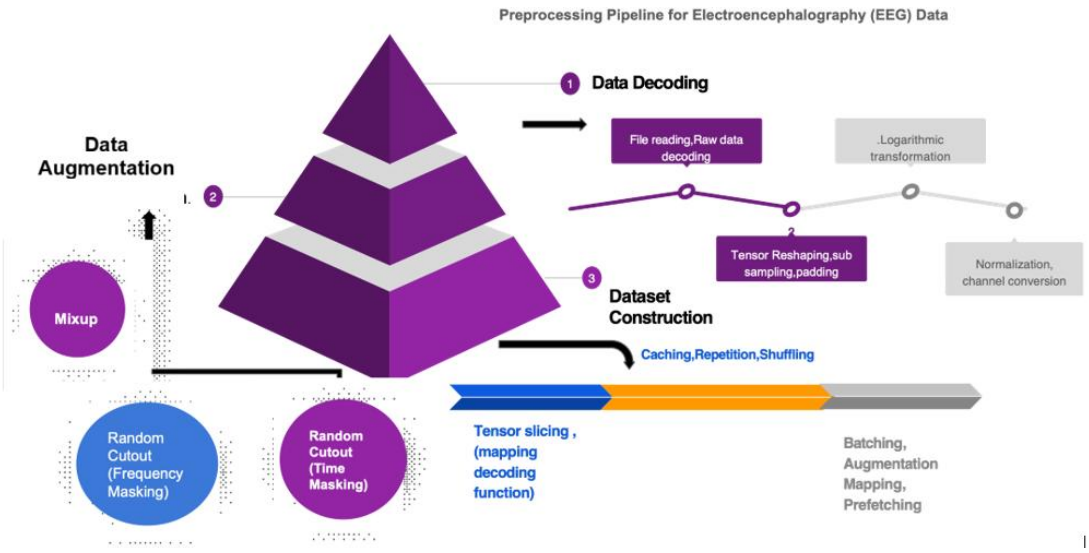

Figure 1. Preprocessing pipeline for EEG data

A logarithmic transformation standardizes high-amplitude variations, smoothing the signal distribution and improving feature interpretability. The EEG data is then reshaped, sub-sampled, and padded reshaping aligns data dimensions with model input, sub-sampling reduces redundancy while preserving temporal information, and padding ensures uniform sequence lengths for batch processing. Normalization and channel conversion ensure consistent scaling across EEG channels. To address limited labeled data, data augmentation is applied: Mixup generates synthetic samples by blending EEG signals, while random cutout (frequency and time masking) removes frequency bands or temporal segments to simulate artifacts and real-world signal gaps. These methods enhance data diversity and model robustness. Finally, batching, augmentation mapping, and preprocessing streamline computational efficiency, delivering high-quality, diverse EEG inputs that enable accurate and robust classification within the DL framework. The DL framework shown in Figure 2.

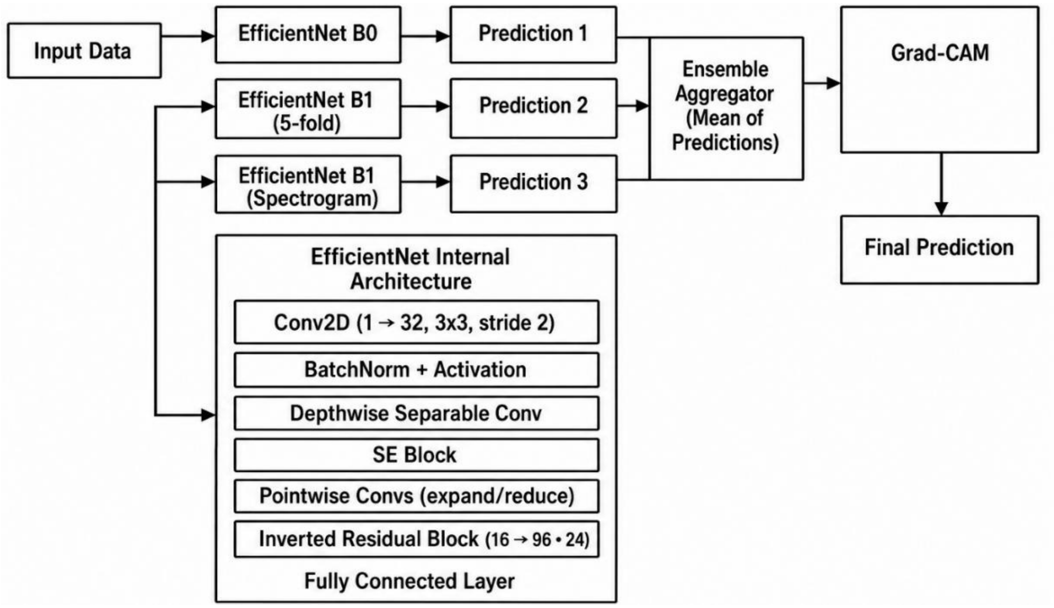

Figure 2. Framework for classification using EfficientNet and Grad-CAM

The input data is fed into three separate EfficientNet-based models: EfficientNet B0, EfficientNet B1 with 5-fold cross-validation, and EfficientNet B1 trained specifically on spectrogram representations of EEG signals. Each model independently produces a prediction, capturing different perspectives of the data. These predictions are combined through an ensemble aggregator that computes the mean of the predictions, improving overall model robustness and reducing individual model bias.

To boost the model interpretability, Grad-CAM is applied to envisagee important regions of the EEG input that influenced the classification, aiding clinical understanding and trust in the automated decision. The internal EfficientNet architecture incorporates several efficient and powerful components: an initial Conv2D layer performs feature extraction, followed by batch normalization and activation functions. Depthwise separable convolutions reduce computational cost while preserving feature quality. Squeeze-and-excitation (SE) blocks recalibrate channel-wise feature responses, enhancing representational power. Pointwise convolutions expand or reduce feature dimensions, and inverted residual blocks enable efficient feature reuse. Finally, a fully connected layer outputs the prediction.

Each block develops the feature dimensions, applies depth wise convolution for spatial feature extraction, and ventures them back to lower dimensions, preserving efficiency while retaining representational power. These weighted blocks progressively capture complex EEG patterns across multiple receptive fields. Following feature extraction, a 1×1 convolution projection layer rises the number of channels to a high-dimensional embedding space, formulating the features for classification. An adaptive average pooling layer then decreases each feature map to a single representative value, guaranteeing that the model is invariant to input size variations. The pooled features are trodden into a one-dimensional vector, which is passed through fully connected layers with dropout.

3. HMS HARMFUL BRAIN ACTIVITY CLASSIFICATION DATASET

All imaging data used in this study were sourced and augmented from the harmful brain activity classification (HBAC) dataset, made available by Harvard Medical School under the CC BY-NC 4.0 license [29]. This dataset includes samples corresponding to six categories: seizure (SZ), generalized periodic discharges (GPD), lateralized periodic discharges (LPD), lateralized rhythmic delta activity (LRDA), and generalized rhythmic delta activity (GRDA). Expert annotators analyzed 50-second EEG segments, aligning them with spectrograms that depicted 10-second windows per minute, focused on the key 10-second interval used for annotation. Overlapping segments were merged to create a cohesive dataset. Each spectrogram image was then resized to 400×300 pixels and paired with its corresponding EEG recording. The dataset was divided into five folds for training, validation, and testing, ensuring no overlap between data splits to eliminate any risk of data leakage and to maintain the integrity of the evaluation process.

4. RESULTS AND DISCUSSION

The interpretability of the proposed model is established through feature importance analysis and LRP, which provide clinicians with meaningful insights into the model's decision-making process. To quantitatively evaluate model performance, the Kullback–Leibler (KL) divergence metric was used to measure the discrepancy across the observed and calculated probability distributions. The model achieved a KL divergence score of 0.30 on the test dataset, indicating strong alignment between predictions and ground truth. Figure 3 elucidates the spectrogram categorized as a seizure due to higher signal strength in the top 50 units. This elevated signal appears consistently across four different timespans. It indicates localized bursts of intense electrical activity in specific EEG segments. Sharp and pronounced spikes reflect strong neuronal firing linked to seizures. These spikes are significantly stronger than the normal background brain waves. Buried waves, representing regular brain activity, are present but masked. The dominant feature remains the abnormal spikes during the seizure event. Therefore, the signal clearly indicates seizure activity throughout the recording.

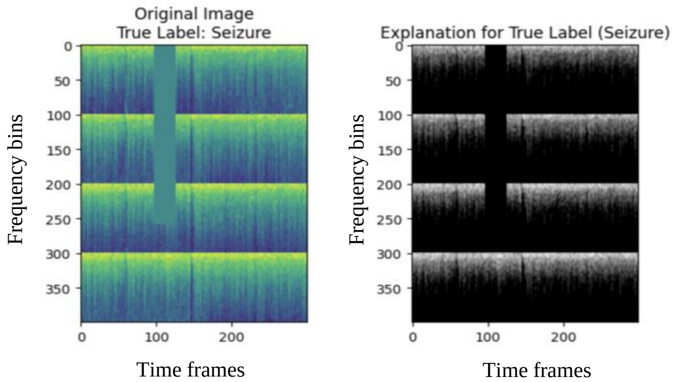

Figure 3. Recorded seizure signal and its explainable heat map showing seizure patterns

Generalised periodic epileptiform discharges (GPEDs) are episodic EEG patterns with bisynchronous sharp wave complexes spiking every 0.5 to 4 seconds. Common in critically ill patients with metabolic or postanoxic encephalopathies, status epilepticus, or progressive CNS diseases (Creutzfeldt–Jacob disease), they indicate poor prognosis. GPEDs are abnormal, repetitive discharges across large brain areas, reflecting underlying brain dysfunction. Key features include high-amplitude, synchronous waves in both hemispheres. Signal strength analysis highlights these bursts, distinguishing them from background EEG. Localized GPD activity appears in the top 25 signal strength units per capture, as shown in Figure 4, confirming the GPD classification.

The categorization of an EEG pattern as a GPD indicates poor neurological activity linked to underlying brain dysfunction. GPDs are often seen in comatose patients, reflecting compromised neural function and uncontrolled neural signals. Their presence signals altered brain function and neurological deterioration, identified by localized bursts of electrical activity, temporal dynamics, comparison with background activity, and clinical correlation. LRDA refers to rhythmic delta waves minimized to one hemisphere, indicating focal brain dysfunction in various neurological disorders. Figure 5 shows the spectrogram classified as LRDA, with abnormal signals appearing early (from 10–15 units), suggesting a persistent abnormal pattern. LRDA commonly occurs in critically ill patients and often progresses into acute seizures, shown by increasing frequency and intensity of abnormal spikes.

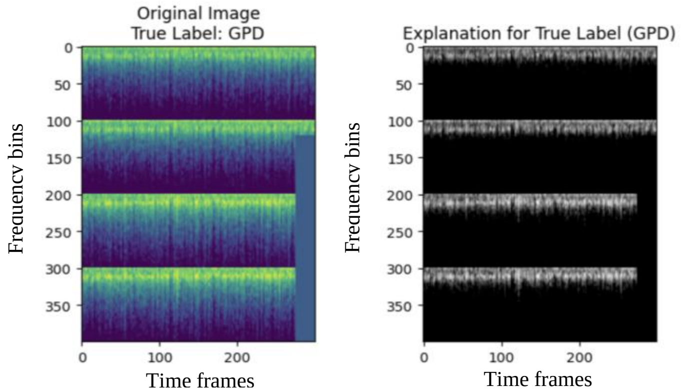

Figure 4. Recorded GPD signal and it's explainable heat map showing GPD patterns

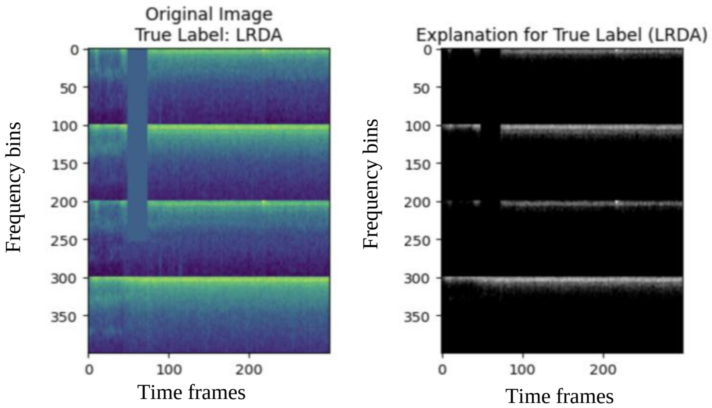

Figure 5. Recorded LRDA signal and it's explainable heat map showing LRDA patterns

This progression signals neurological deterioration and seizure onset, frequently observed in ICU patients due to pre-existing conditions or traumatic insults. LRDA thus serves as an important marker of severe neurological decline. These findings reveal the spatial-temporal features of LRDA in EEG, aiding the interpretation of abnormal brain activity in critically ill patients. When higher frequency synchronization or desynchronization occurs without sharp activity changes, the EEG signals appear continuous and uniform in frequency and amplitude, without signs of seizures, periodic discharges, rhythmic delta activity, or other pathological patterns. The spectrogram in Figure 6 is categorized as "Other," as it lacks sharp frequency peaks and does not show any damaging brain activity, suggesting the signals represent normal physiological fluctuations or benign variations unrelated to disease. GRDA refers to rhythmic delta waves involving both hemispheres, often indicating brain dysfunction. In Figure 7, the spectrogram is classified as GRDA because abnormal signals emerge early (from 15–20 units) and display significantly higher frequencies than LRDA. GRDA arises from metabolic disturbances, cerebral lesions, physical damage, or insufficient blood/oxygen supply. These abnormal patterns are evident from the recording's start, affecting both hemispheres, and show faster delta rhythms than LRDA, indicating a more aggressive dysfunction. GRDA is commonly linked to metabolic encephalopathy and structural abnormalities (traumatic brain injury, stroke). The classification is supported by early abnormal bilateral signals, higher-frequency spikes than LRDA, and metabolic changes related to cerebral lesions. These insights highlight the clinical relevance of GRDA in detecting and interpreting pathological brain activity in EEG recordings. The Table 1 provides a comparative analysis of three ML models—EfficientNet Ensemble with Grad-CAM, CNN, and SVM—across multiple key performance metrics for EEG classification, particularly in detecting harmful brain activities in critically ill patients. In terms of accuracy, CNN leads slightly (99.04%–99.96%) compared to EfficientNet ensemble (98.8%–99.6%), while SVM lags significantly (72%–96%). For explainability, EfficientNet ensemble achieves the highest (85%–95%) due to integrated Grad-CAM visualizations, providing clear heatmaps of relevant EEG regions. CNN offers moderate explainability (50%–70%), typically using post-hoc methods. SVM has low explainability (10%–30%) due to reliance on manual feature engineering without clear visual interpretation. Regarding computational efficiency, SVM is the fastest (~0.1–1 ms/sample) but works best with small datasets. CNN offers balanced efficiency (~5–15 ms/sample), while EfficientNet Ensemble is slower (~15–50 ms/sample) because of model complexity. For generalization on small data, EfficientNet Ensemble (85%–95% accuracy) outperforms CNN (70%–85%) and SVM (60%–75%), largely thanks to augmentation strategies. In handling complex EEG patterns (Seizure, GPD, LRDA, and GRDA), EfficientNet Ensemble leads (90%–98%), followed by CNN (80%–95%) and SVM (65%–85%). In robustness to artifacts, EfficientNet Ensemble retains 85%–95% accuracy, CNN 70%–85%, and SVM 40%–70%. Finally, for interpretation of predictions, only EfficientNet provides clear Grad-CAM visualizations with high clinician agreement (80%–95%), while CNN and SVM have much lower interpretability.

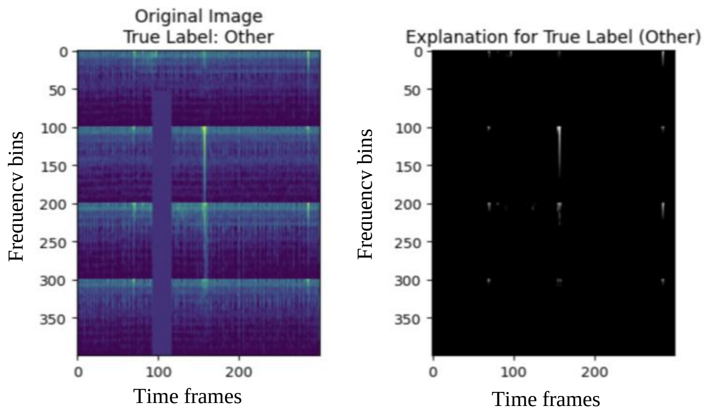

Figure 6. Recorded “other” signal and it’s explainable heat map showing no patterns pertaining to the HBAC classes

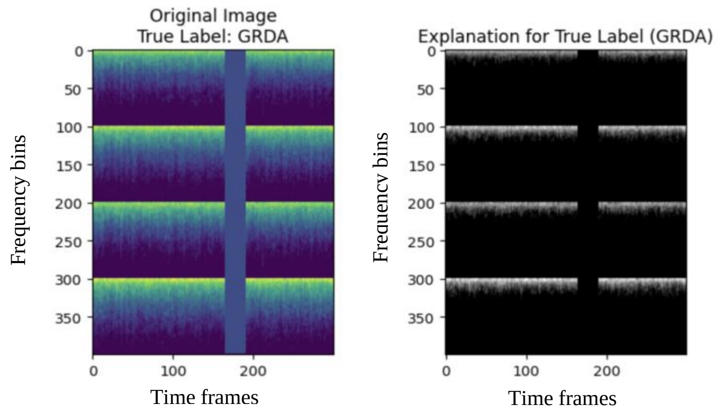

Figure 7. Recorded GRDA signal and its explainable heat map showing GRDA patterns

Table 1. Comparative analysis of EfficientNet Ensemble, CNN, and SVM

Feature/metric	EfficientNet ensemble+Grad-CAM	CNN	SVM
Accuracy	98.8%-99.6%	99.04%–99.96%	72%–96%
Explainability	85%-95%	50%-70%	10%-30%
Computational efficiency	~15–50 ms inference time per sample	~5–15 ms inference time per sample	~0.1–1 ms inference time per sample (but fast inference with small time sample)
Generalization on small data	85%–95% accuracy	70%–85% accuracy	60%–75%
Handling complex patterns (seizure, GPD, LRDA, and GRDA)	90%–98% accuracy/F1-score	80%–95% accuracy/F1-score	65%–85% accuracy/F1-score
Robustness to artifacts	85%–95% accuracy retention	70%–85% accuracy retention	40%–70% accuracy retention
Interpretation of predictions	Visualized via Grad-CAM; 80%–95% clinician agreement/localization accuracy	50%–70% clinician agreement/localization accuracy. Without explicit explanation mechanisms, interpretation is harder	20%–40% clinician agreement/relevance to clinical features; typically relies on feature weights or importance scores, which are harder to map

The confusion matrix shown in Figure 8 illustrates the model's performance in classifying five EEG patterns: Seizure, GPD, LRDA, and GRDA. For Seizure, 185 samples were correctly classified, with only 3 misclassified as GPD. In the GPD category, 180 samples were accurately predicted, while 2 were misclassified as Seizure and 1 as LRDA. The LRDA class had 175 correct predictions, with 2 misclassified as GPD and 1 as GRDA. For GRDA, 170 samples were correctly identified, with a single misclassification as LRDA. The Other category achieved perfect classification, with all 260 samples correctly identified. The results indicate high model accuracy and strong ability to differentiate complex EEG patterns, as the majority of predictions fall on the diagonal, and misclassifications are minimal.

Figure 8. Confusion matrix for EEG pattern classification

5. CONCLUSION

This study bridges the gap between DL and clinical interpretability by providing explainable EEG classification using an ensemble model. Future work will focus on clinical validation and further refinement of interpretability techniques. This novel approach bridges the gap between automated DNN based methods and clinical interpretability, allowing state-of-art performance on a publicly available EEG dataset. The method demonstrates its potential to overcome the limitations of current labor-intensive workflows, improve diagnostic accuracy, and support clinical decision making. This contribution represents a significant step forward integrating DL model into practical neurological applications, enabling timely interventions and enhancing patient care. From this study, in future this work can be extended to validate the proposed DNN model on larger, diverse, multi-center EEG data sets to ensure its robustness and generalizability across different populations and recording conditions. It is also recommended to incorporate the user feedback from clinicians to enhance the serviceability of interpretability features.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Manikandan Arunachalam	✓				✓					✓	✓	✓	✓	
Sanjay Thiyagarajan		✓	✓	✓	✓		✓	✓	✓				✓	
Nagandla Chirudeep			✓	✓	✓	✓	✓			✓				

C	: **C**onceptualization	I	: **I**nvestigation	Vi	: **Vi**sualization
M	: **M**ethodology	R	: **R**esources	Su	: **Su**pervision
So	: **So**ftware	D	: **D**ata Curation	P	: **P**roject administration
Va	: **Va**lidation	O	: Writing - **O**riginal Draft	Fu	: **Fu**nding acquisition
Fo	: **Fo**rmal analysis	E	: Writing - Review & **E**diting		

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] S. Singla *et al.*, "Detecting Seizures and Epileptiform Abnormalities in Acute Brain Injury," *Current Neurology and Neuroscience Reports*, vol. 20, no. 9, p. 42, Sep. 2020, doi: 10.1007/s11910-020-01060-4.

[2] J. Kang, Y. J. Jeong, and J. Hong, "The effect of postintensive care syndrome on the quality of life of intensive care unit survivors: A secondary analysis," *Australian Critical Care*, vol. 34, no. 3, pp. 246–253, May 2021, doi: 10.1016/j.aucc.2020.08.006.

[3] P. Hoxhaj *et al.*, "Investigating the Impact of Epilepsy on Cognitive Function: A Narrative Review," *Cureus*, vol. 15, no. 6, Jun. 2023, doi: 10.7759/cureus.41223.

[4] V. Bhandari and D. H. Manjaiah, "Improved ensemble learning model with optimal feature selection for automated epileptic seizure detection," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, vol. 11, no. 2, pp. 135–165, Mar. 2023, doi: 10.1080/21681163.2022.2058616.

[5] A. Marey *et al.*, "Explainability, transparency and black box challenges of AI in radiology: impact on patient care in cardiovascular radiology," *Egyptian Journal of Radiology and Nuclear Medicine*, vol. 55, no. 1, pp. 1–14, Sep. 2024, doi: 10.1186/s43055-024-01356-2.

[6] A. A. E. Shoka, M. M. Dessouky, A. El-Sayed, and E. E. D. Hemdan, "EEG seizure detection: concepts, techniques, challenges, and future trends," *Multimedia Tools and Applications*, vol. 82, no. 27, pp. 42021–42051, Nov. 2023, doi: 10.1007/s11042-023-15052-2.

[7] T. M. Ganguly *et al.*, "Seizure Detection in Continuous Inpatient EEG: A Comparison of Human vs Automated Review," *Neurology*, vol. 98, no. 22, pp. E2224–E2232, May 2022, doi: 10.1212/WNL.0000000000200267.

[8] X. Wang, F. Yang, B. Chen, and W. Jiang, "Non-convulsive seizures and non-convulsive status epilepticus in neuro-intensive care unit," *Acta Neurologica Scandinavica*, vol. 146, no. 6, pp. 752–760, Dec. 2022, doi: 10.1111/ane.13718.

[9] V. Murugaiyan, K. S. K. Jebaraj, and S. Kalaiselvan, "Current trends in neurocritical care: a comprehensive review," *Journal of Neurocritical Care*, vol. 18, no. 1, pp. 1–15, Jun. 2025, doi: 10.18700/jnc.240034.

[10] D. Dhanush, V. T. Prakasham, S. A. Saathappan, G. Harissh, and M. Ganesan, "Integrated Critical Care Monitoring System for Ventilated Patients: ECG Analysis and Air Flow Detection," in *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, IEEE, Dec. 2023, pp. 917–924, doi: 10.1109/UPCON59197.2023.10434563.

[11] G. S. Jebaraj and K. Elango, "A Comprehensive Review of EEG-Based Seizure Detection Techniques," *IEEE Access*, vol. 13, pp. 103531–103564, 2025, doi: 10.1109/ACCESS.2025.3578991.

[12] S. Sharma, M. Nunes, and A. Alkhachroum, "Adult Critical Care Electroencephalography Monitoring for Seizures: A Narrative Review," *Frontiers in Neurology*, vol. 13, Jul. 2022, doi: 10.3389/fneur.2022.951286.

[13] S. Ghildiyal, R. Garg, S. Dhanik, A. Sarkar, G. Sharma, and P. S. Kharayat, "A Comprehensive Review Analysis of Supervised Machine Learning Techniques," in *2024 5th International Conference on Smart Electronics and Communication (ICOSEC)*, IEEE, Sep. 2024, pp. 925–930, doi: 10.1109/ICOSEC61587.2024.10722516.

[14] Sudimanto, I. H. Kartowisastro, A. Trisetyarso, and W. Budiharto, "Pain Classification Using Discrete Wavelet Transform Feature Extraction and Machine Learning Techniques," *IEEE Access*, vol. 13, pp. 45912–45922, 2025, doi: 10.1109/ACCESS.2025.3545869.

[15] O. Olawuyi, "Detection and Classification of Prostate Cancer Using Lightweight MobileNet," in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 2025, pp. 15–33, doi: 10.1007/978-3-031-94442-0_2.

[16] H. F. Atlam, G. E. Aderibigbe, and M. S. Nadeem, "Effective Epileptic Seizure Detection with Hybrid Feature Selection and SMOTE-Based Data Balancing Using SVM Classifier," *Applied Sciences*, vol. 15, no. 9, pp. 1–26, Apr. 2025, doi: 10.3390/app15094690.

[17] L. A. Atmuri *et al.*, "Fast Iterative Filtering based approach for Seizure Detection in EEG Signals," in *2024 11th International Conference on Advances in Computing and Communications (ICACC)*, IEEE, Nov. 2024, pp. 1–6, doi: 10.1109/ICACC63692.2024.10845336.

[18] R. Devisetty *et al.*, "Localizing epileptogenic network from SEEG using non-linear correlation, mutual information and graph theory analysis," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 236, no. 12, pp. 1783–1796, Dec. 2022, doi: 10.1177/09544119221134991.

[19] A. Rauniyar, B. M. Reddy, and D. Remya, "Stratifying Prostate Cancer: A Comprehensive Framework for Staging and Treatment Planning Classification," *Procedia Computer Science*, vol. 259, pp. 356–365, 2025, doi: 10.1016/j.procs.2025.03.337.

[20] K. Gopikrishna, N. R. Niranjan, S. Maurya, V. G. U. Krishnan, and S. Surendran, "Automated Classification and Size Estimation of Fetal Ventriculomegaly from MRI Images: A Comparative Study of Deep Learning Segmentation Approaches," *Procedia Computer Science*, vol. 233, pp. 743–752, 2024, doi: 10.1016/j.procs.2024.03.263.

[21] A. Tocchetti and M. Brambilla, "The Role of Human Knowledge in Explainable AI," *Data*, vol. 7, no. 7, pp. 1–20, Jul. 2022, doi: 10.3390/data7070093.

[22] F. A. Khan, Z. Umar, A. Jolfaei, and M. Tariq, "Explainable AI for epileptic seizure detection in Internet of Medical Things," *Digital Communications and Networks*, vol. 11, no. 3, pp. 587–593, Jun. 2025, doi: 10.1016/j.dcan.2024.08.013.

[23] A. S. Alharthi *et al.*, "Explainable AI for Sensor Signal Interpretation to Revolutionize Human Health Monitoring: A Review," *IEEE Access*, vol. 13, pp. 115990–116024, 2025, doi: 10.1109/ACCESS.2025.3585764.

[24] A. M. Kabeer, R. K. Megalingam, and K. M. Sakthiprasad, "Brain–Computer Interfaces for Mobility Assistance: A Comparative Analysis of EEG-Based Classification Models for Brain-Controlled Wheelchairs," in *Smart Innovation, Systems and Technologies*, 2024, pp. 245–256, doi: 10.1007/978-981-97-6810-3_20.

[25] A. Manikandan and T. Sanjay, "Generative adversarial networks: A game changer - GAN for machine learning and IoT applications," in *Handbook of Research on Machine Learning-Enabled IoT for Smart Applications Across Industries*, 2023, pp. 67–83, doi: 10.4018/978-1-6684-8785-3.ch004.

[26] I. Andrade, C. Teixeira, and M. Pinto, "On the performance of seizure prediction machine learning methods across different databases: the sample and alarm-based perspectives," *Frontiers in Neuroscience*, vol. 18, Jul. 2024, doi: 10.3389/fnins.2024.1417748.

[27] C. Wang, X. Yuan, and W. Jing, "Artificial intelligence in electroencephalography analysis for epilepsy diagnosis and management," *Frontiers in Neurology*, vol. 16, Aug. 2025, doi: 10.3389/fneur.2025.1615120.

[28] A. Gupta, R. Chauhan, G. Saravanan, and A. Shreekumar, "Improving sepsis prediction in intensive care with SepsisAI: A clinical decision support system with a focus on minimizing false alarms," *PLOS Digital Health*, vol. 3, no. 8, pp. 1–16, Aug. 2024, doi: 10.1371/journal.pdig.0000569.

[29] J. Jing *et al.*, "HMS - Harmful Brain Activity Classification," *Kaggle*, 2024, Online [Available]: https://kaggle.com/competitions/hms-harmful-brain-activity-classification. (Accessed: Mar. 03, 2024).

BIOGRAPHIES OF AUTHORS

Dr. Manikandan Arunachalam completed his B.E. in Electronics and Communication Engineering from Madurai Kamaraj University in 2004, Master of Engineering in Communication Systems from Anna University in 2006. He received his Ph.D. from Anna University at 2018. He has 18 years of teaching experience and guided many UG and PG Projects. He is a Life member of IETE and ISTE. He has published papers in 15 international journals, 20 international conferences and 6 national conferences. Currently he is working in M2M communication and IoT. He is currently associated with Amrita School of Engineering, Amrita Vishwa Vidhyapeetham, Coimbatore, Tamilnadu India. He can be contacted at email: a_manikandan@cb.amrita.edu or mani_ac18@yahoo.co.in.

Sanjay Thiyagarajan is currently pursuing Master's degree in Artificial Intelligence and Machine Learning at Birla Institute of Technology and Science (BITS) Pilani. His research interests include neural networks, generative adversarial networks, machine learning algorithms, and agentic orchestration. He can be contacted at email: sanjaythiyagarajan@gmail.com.

Nagandla Chirudeep is currently pursuing a bachelor's degree in Electronics and Computer Engineering at Amrita Vishwa Vidyapeetham, Amritapuri Campus, Kollam, Kerala. He has worked on various electronics and software-based projects, including datalogger to keep update on landslide prone areas. His research interests include machine learning, embedded systems, digital signal processing, and medical AI applications. He can be contacted at email: chirudeepnagandla308@gmail.com.