Robust Seizure Prediction Model Using EEG Signal for Temporal Lobe Epilepsy Leveraging Deep Learning and Continual Learning

Bikash Chandra Sahoo[1], Shruti Mishra[2], Member, IEEE , Sandeep Kumar Satapathy[3], Member, IEEE, Sachin Kumar[4], Senior Member, IEEE

[1]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India- 600127
[2,3]Centre for Advanced Data Science, Vellore Institute of Technology, Chennai, Tamil Nadu, India- 600127
[4]Akian College of Science and Engineering, American University of Armenia, Yerevan, 0019, Armenia

Corresponding author: Shruti Mishra (e-mail: shrutim2129@gmail.com).

"This work was supported by Vellore Institute of Technology, India"

ABSTRACT Epilepsy is characterized by abnormal electrical activity in the brain that affects millions worldwide and poses significant challenges to patients and healthcare providers. Seizure prediction has emerged as a pivotal area of research in epilepsy management aiming to mitigate the unpredictability of seizures and enhance patient quality of life. This research focuses on leveraging memory-based learning approaches including transfer learning and recurrent neural networks (RNNs) for seizure prediction. The dataset comprises time-domain signals embedded with noise necessitating preprocessing techniques such as filtering for optimal utilization. By employing deep learning methodologies, the work seeks to classify seizure in temporal lobe epilepsy from the EEG signals to improve seizure prediction accuracy. With deep learning ability to capture temporal dependencies in sequential data, they offer promising avenues for modeling the temporal dynamics of epileptic seizures. Among the performed architecture and their merits when compared to other, "Googlenet" proved to be the best architecture for training the model by achieving the accuracy of 97.5%. Through continual learning strategies, the model's training process adapts dynamically enabling it to incorporate new information while retaining knowledge from previous tasks. We used "Elastic Weight Consolidation (EWC)" as the continual learning technique which enforces adaptability and robustness of the trained model. With a custom learning rate scheduler that adjust the learning rate during the training based on the current epoch in the EWC technique we enhanced the accuracy to 98.7% which is pretty much promising result. The model improved by 1.2% over the baseline GoogLeNet (97.5%) to reach an accuracy of 98.7% by including Elastic Weight Consolidation (EWC) for continuous learning. When paired with a customized learning rate scheduler, EWC accelerates convergence and improves model adaptability by eliminating catastrophic forgetting.

INDEX TERMS Epilepsy; EEG; Deep Learning; Transfer Learning; Continual Learning; EWC

I. INTRODUCTION

The evolution of technology with its involvement in other domain has greatly impacted human lives. One such domain where technology plays a vital role is the medical field where the diagnostic and predictive analysis has become way easier. With its integration with clinical and other applications, it now fosters better results in improvising healthcare. The contribution of machine learning on analyzing and predicting epilepsy is highly commendable which helps doctors to a great extent.

Epilepsy is a condition where people have repeated and unpredictable seizures and affects around 65 million individuals worldwide. To predict seizures, doctors usually rely on watching symptoms and using EEG tests that monitor brain activity by placing electrodes on the scalp. Unfortunately, these methods sometimes cannot predict seizures accurately or in time, leaving patients at risk of sudden attacks. [1] Living with epilepsy can greatly affect a person's life quality exposing them to dangers of injuries and social stigma. There are different types of seizure depending on the region of brain where it faces disturbance. The Seizure occurring in the temporal lobe and the frontal lobe [2] are the common scenario. The temporal lobe epilepsy (TLE) results in the absence seizure which is also the case that happens to an individual at least once in the lifetime. For measuring epilepsy, electrodes with certain configuration are placed on the scalp of head. The standard configuration which is 10-20 is followed in order to maintain particular distance among the electrodes. Depending on the electrode's placement, whether it is placed above the brain cell or in touch with brain cell, they are classified into two namely subcutaneous and intercranial measurement.

While measuring brain activity in the form of electric signals measured in microvolts, they are always interfered by different types of noises captured from body movement as well as the environment. [3] Due to many electrodes (generally 21) measurement that to each electrode measuring signal of very small instances in order to track minute but strong fluctuation there are many features in the data used for the analysis. The features either need to be reduced or particular feature contributing the most in the analysis need to be considered for implementing a particular phase of processing.

As mentioned, the EEG data is always compromised by certain artifacts (noise) which should always be pre-processed to further make it available for training a model. [4] To de-noise an EEG signal we require filters which are of different types and are used based on the goal. For eliminating noise signal, we are availed with integral filters like linear filters, adaptive filters, wavelet filters etc. Since EEG is a time series data and filters are governed by cut-off frequency, the signal is always converted to frequency domain thus further processed by filter for de-noising. The performance of filter's de-noising is dependent on its signal to noise ratio (SNR) which determines the extent of noise removal from the signal. The higher the SNR value better will be the signal quality.

Analyzing epilepsy technically faces many challenges because epilepsy is complex and the methods we have are limited. One big challenge is getting useful information from EEG signals which are commonly used to predict epilepsy. [5] Understanding EEG data needs special knowledge and personal opinions can affect the analysis making it more complicated. Another challenge in analyzing epilepsy technically is that seizures change a lot. They happen in different ways and at different times, which makes it tough to make models that can catch all these changes accurately. Old methods might not be able to keep up with these changes or the differences between patients over time.

The seizure patterns could be observed in short instances of time which was captured for a long time (in hours). In order to capture the period of exact seizure, the whole waveform is divided into windows or wavelets [6]. This not only helps in analyzing the pattern well but are also helpful in processing the long waveform in batches.

Machine Learning has proved its worth in fetching insights and predicting certain task while dealing with time series data[7-8]. Advances in machine learning and brain-focused computing reignited interest in predicting epilepsy. Researchers began exploring more complex algorithms like artificial neural networks and support vector machines to understand the intricate timings in EEG data.

With the introduction of deep learning algorithms [9], the way to classify certain problem has improved to a large extent. The neural network has the biggest advantage over machine learning algorithms which is to identify non-linearity among the features. This gives an edge over others as the data consisting of time series are generally dynamic in nature. Finding relation from such series and henceforth training it becomes an easy task. For training an EEG data basic ANN and CNN model has proved to be very efficient. The way data is growing and changing in a rapid pace it is very important to use technique where it can remember the training and pattern [10]. This is where memory-based learning came in clutch. Architectures like Recurrent Neural Network (RNN) and transfer learning are very useful in acquainting the model with the insights it was already trained upon.

Seizure dynamics can change over time due to the factors like medication adjustments, lifestyle changes or disease progression. [11] This means prediction models must adapt to these changes and shifts in seizure patterns. Adapting requires constant refinement and updating of models which raises logistical and computational challenges for real-time prediction systems. One such method is continual learning which helps in automating a model. This method not only helps in adapting to the knowledge which was learnt already but also helps in improving the accuracy of the model, thus enforcing the robustness of the model. Our method is unique in that it incorporates Elastic Weight Consolidation (EWC) for continuous learning, addressing catastrophic forgetting, a significant drawback of conventional deep learning models. EWC prevents this by maintaining significant weights from past jobs, whereas traditional models often lose previously learnt information when trained on fresh data. This greatly improves the model's long-term adaptability and robustness, which makes it more appropriate for dynamic, real-world EEG data where patterns might change over time. Our model improves seizure prediction accuracy and reliability by utilizing a unique learning rate scheduler in conjunction with GoogleNet's optimal performance.

The continual learning adopts several techniques to avoid catastrophic forgetting. When used in combination with deep learning, it eliminates the factor of ignoring patterns and other insights learnt during the whole process of training. [12] The most basic technique for implementing continual learning is rehearsal-based learning. In this, after the training of the model, the new data is mixed with the old ones to avoid forgetting. The approach is simple but not robust enough to retain the knowledge gained from the basic training. This is where new and more techniques were introduced which were far more robust but are highly complex to implement.

The continual learning may prove a significant role in analyzing epilepsy but it still lacks several other data which needs to be considered for more precise prediction especially in ictal phase. [13] The multi-modal approach may break this barrier of unpredictability and dynamicity of seizure. The multi-modal approach makes use of different data which contributes during the seizure. By analyzing multiple data of an epileptic individual, it would create more connection that triggers an epilepsy thus could help doctors diagnose the patient beforehand.

Though machine learning and other techniques have come so far in forecasting medical issues like epilepsy, [14] we cannot ignore the fact that the amount of data is growing in a rapid manner and the factor affecting certain issues are also changing based on an individual's routine. So, in order to deal with the huge amount of data, big data need to be brought in to the picture while [15] simultaneously incorporating explainable AI and generative AI. By implementing these actions and validating the predictions with medical experts, one could develop a highly precise and sustainable model for predicting epilepsy. Also, three main differences set our model apart from other deep learning-based seizure detection methods: (1) it incorporates Elastic Weight Consolidation (EWC) to facilitate continuous learning and adaptability to new data; (2) it finds GoogleNet to be the best architecture for extracting multiscale EEG features; and (3) it makes use of a unique learning rate scheduler to increase training accuracy and efficiency. Together, these improvements raise the model's prediction performance and robustness above and beyond that of conventional techniques.

II. LITERATURE SURVEY

The epilepsy has always been a dynamic domain which has piqued the interest of many analysts as well the researchers in order to draw accurate analysis for diagnostics. Following are some of the researches done for reference from reliable sources like IEEE, PubMed etc.

One of the types of research work carried out by Hartmann et al. [16] used 102 twenty second seizure onset EEG from 48 consecutive epilepsy as the dataset. The work is accompanied by a proposed model called PureEEG which is fully automatic EEG interference or noise removal method. In this, EEG recordings with noise (artifacts) of various types were targeted and was given as an input. The response showed improvement in EEG quality post-processing with majority of EEG epochs showing major or minor improvements and no major attenuation of significant EEG patterns. The work also involved validation by two independent reviewers in order to prove the performance of algorithm in various conditions. The EEG pattern was more readable and easier to fetch insights. Hartmann et al. [17] conducted a study using the Temple University Hospital Seizure Detection Corpus as their dataset utilizing a deep neural network specifically designed for two-channel seizure detection. The model processed EEG recordings that were captured using the 10-20 electrode system and were annotated to indicate seizure events. This advanced model architecture comprised several neural network elements including convolutional layers and a multilayer perceptron (MLP) classifier which are integral for analyzing complex patterns in EEG signals. The study's outcome demonstrated a mean sensitivity of 88.8% in detecting seizures across patients with a daily false positive rate of 12.9. These results underscore the model's robustness and its potential for real-world application in monitoring and diagnosing epilepsy through minimalistic two-channel EEG systems. The promising performance also suggests that this approach could be a valuable tool in clinical settings, offering a reliable method for continuous epilepsy monitoring without extensive hardware.

Remvig et al. [18] utilized ultra long-term subcutaneous EEG data from nine individuals with epilepsy. They deployed a sophisticated deep neural network constructed using TensorFlow which featured multiple layers specifically tailored to process the subcutaneous EEG recordings effectively. This innovative approach enabled the automatic detection of seizures, which was subsequently validated through comparison with expert annotations by human clinicians. The results of this study showed a sensitivity rate of 86% in accurately identifying seizures alongside a low false detection rate of 2.4 per 24 hours. These findings affirm the clinical utility of this semi-automatic review process for sqEEG-based seizure detection. Additionally, the low rate of false detections and high sensitivity highlight the potential of this technology to provide reliable and continuous monitoring for epilepsy patients potentially improving patient outcomes and quality of life by allowing for better-managed care and early intervention strategies.

Furbass et al. [19] studied EEG, EMG and ECG recordings from 92 epileptic patients using a computer algorithm designed to process these signals. Their innovative method enabled the automatic detection of seizures by leveraging long-term monitoring of EEG, EMG and ECG data. The study achieved an overall detection sensitivity of 86%, coupled with an average false detection rate of 16.5 per 24 hours. Notably, the sensitivity of the detection varied significantly across different types of epilepsy and specific seizure manifestations. This variability in performance underlines the complex nature of epilepsy, which can manifest differently in patients, affecting the algorithm's efficacy.

Bhattacherjee et al. [20] work was completely bootstrapped and gathered real-time data from epileptic seizure patients using a low-cost and portable device. They utilized a Dual Tree-Complex Wavelet Transform for processing the data which was then analyzed using an SVM classifier. This sophisticated approach allowed for the effective scaling and filtering of EEG data enabling accurate classification between pre-ictal (before seizure) and inter-ictal (between seizures) states via a hyperplane established by the

SVM. The study reported an impressive accuracy rate of 85.9% with the capability of predicting seizures up to 60 minutes before their onset. These results underline the significant potential of this technology for early seizure detection which could be crucial in managing epilepsy more effectively.

Masum et al. [21] investigated an EEG dataset from the UCI machine learning repository by applying a range of machine learning models such as SVM, random forest, decision tree and deep neural networks. These models were tasked with analyzing both the original EEG data and seven additional datasets each comprising various combinations of principal components to enhance feature detection and model accuracy. The study meticulously calculated the weighted F-beta score to assess the balance between precision and recall and applied the Wilcoxon Rank Sum test to determine the statistical significance of the results. The findings indicated that each model SVM, RF, DT and DNN excelled on different dataset configurations, with the DNN model showing the highest statistical significance against both the original and transformed datasets. This suggests that the deep neural network was particularly effective at leveraging complex, high-dimensional data for seizure detection. The success of the DNN underscores its potential as a robust tool in the field of epilepsy research offering significant implications for future predictive analytics and personalized medicine approaches in neurological disorders.

Sharma and Arora [22] employed the CHB-MIT EEG dataset to evaluate the efficacy of various machine learning classifiers such as SVM, decision trees (DT) and XG Boost. They preprocessed the EEG data using Principal Component Analysis (PCA) to reduce dimensionality and enhance the classifier's ability to detect patterns. A thorough comparative analysis was then conducted, with a particular focus on metrics such as accuracy and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). The results from their study highlighted that XG Boost initially led the performance with an impressive accuracy of 96.6%. However, after optimizing the parameters, the SVM classifier surpassed this achieving a peak accuracy of 98.5%. This improvement underscores the potential of fine-tuning in achieving superior performance in seizure detection algorithms.

Salvador et al. [23] conducted research using the CHB-MIT EEG dataset, applying a bagging tree model to their analysis. They focused on refining EEG data by denoising and extracting key features to enable more effective classification of ictal (during seizure) and pre-ictal (before seizure) states across different EEG channel combinations. Their evaluation strategy emphasized metrics such as accuracy, precision and recall which were assessed not only for various channel combinations but also across individual patient data. Although their initial findings provided valuable insights into the effectiveness of the bagging tree model for seizure prediction the analysis did not extend into a comprehensive evaluation across all possible variables.

Nigam et al. [24] utilized intracranial EEG data provided by the American Epilepsy Society, engaging a diverse array of models such as decision trees, random forest, SVM and regression to analyze the data. The EEG signals were initially preprocessed using FFT band filters to isolate relevant frequencies and feature extraction techniques to enhance the detectable patterns within the data. Their methodology focused on both window-level and segment-level seizure prediction aiming to identify seizures within specific, confined intervals of data as well as over longer continuous segments. Their findings highlighted that the implementation of windowing techniques significantly enhanced prediction accuracy, optimizing the model's ability to accurately detect seizure events. Particularly, the random forest model exhibited robust performance across various patient datasets indicating its effectiveness in handling the complex variations present in intracranial EEG recordings.

Rohan et al. [25] conducted an analysis on an EEG dataset obtained from the UCI machine learning repository applying XG Boost and Artificial Neural Network (ANN) models to evaluate their effectiveness in seizure detection and classification. The dataset comprised EEG recordings from five distinct cases, providing a diverse set of data for the models to process. The primary objective was to not only detect epileptic seizures but also to accurately classify these seizures into one of the five predefined categories. In this study the ANN model demonstrated superior performance compared to the XG Boost model achieving an impressive accuracy rate of 98.2%. This high level of accuracy which surpassed that of previously implemented models highlights the potential of ANNs in handling complex patterns within EEG data effectively.

Seifi et al. [26] conducted a study leveraging ECG signals from 13 patients at the University of Siena utilizing a combination of sophisticated classifiers namely SVM, KNN and Naïve Bayes. These signals, marked by distinct time and frequency domain features, served as the foundation for effectively distinguishing between pre-ictal and interictal states. Of the classifiers tested, Naïve Bayes emerged as the standout performer demonstrating superior efficiency and accuracy over both SVM and KNN. This finding emphasizes the robustness of Naïve Bayes in processing complex, time-sensitive physiological data crucial for predicting epileptic seizures. The success of Naïve Bayes in this study not only reinforces its applicability in ECG-based seizure forecasting but also suggests significant advancements in real-time seizure monitoring technologies.

Qin et al. [27] delved into the effectiveness of integrating CNN (Convolutional Neural Network) with ELM (Extreme Learning Machine) to classify pre-ictal and interictal states from the CHB-MIT EEG dataset. This approach utilized 2D spectrograms, which were transformed from raw EEG time series, serving as the input data. By harnessing the CNN's robust feature extraction capabilities along with the ELM's rapid learning efficiency, the model achieved a remarkably high sensitivity rate of 95.85% and maintained a minimal false prediction rate of 0.045 per hour. This dual-technology methodology not only significantly boosted the precision of seizure predictions but also laid a groundwork for further advancements in EEG analysis. It points towards enhancing real-time seizure monitoring systems which could drastically improve patient management in clinical settings. The integration of CNN and ELM exemplifies a forward-thinking approach to neurological diagnostics, potentially leading to breakthroughs in how seizures are predicted and managed using EEG data.

Wang et al. [28] utilized the Freiburg iEEG database for their investigation implementing a 1D CNN model enhanced by a channel increment strategy. This approach meticulously processed intracranial EEG signals that were segmented into non-overlapping 4-second windows aiming specifically to differentiate between pre-ictal and interictal states. The model demonstrated exceptional sensitivity and consistently low false prediction rates at both the segment and event-based evaluations, markedly surpassing the performance of many earlier studies. This superior performance underscores the effectiveness of the channel increment strategy combined with 1D CNN in identifying fine-grained neural variations that are crucial for accurate seizure prediction. The success of proposed method suggests a significant advancement in the domain of clinical seizure prediction proposing a potent tool for neurologists and clinicians to forecast and manage epilepsy more effectively.

Chen et al. [29] conducted an innovative study using the CHB-MIT EEG dataset to evaluate an Auto Machine Learning (Auto ML) system that combines various machine learning algorithms. This advanced system was specifically designed to classify EEG signals into seizure and non-seizure states. Enhanced by CTGAN (Conditional Generative Adversarial Network) for the generation of synthetic data the Auto ML system achieved an exceptional average accuracy of 99%. This high level of accuracy highlights the effectiveness of utilizing automated machine learning which facilitates adaptive algorithm selection and hyperparameter optimization thus significantly refining model performance. The success of this approach not only underscores the practicality of Auto ML in the field of neurology but also establishes a new standard for future research in automated EEG signal classification.

In a pioneering study by Dissanayake et al. [30], researchers used the CHB-MIT-EEG and Siena-EEG datasets to create a new seizure prediction model that operates independently of the subject using a Graph Neural Network. This model transformed EEG data into Mel-frequency Cepstrum Coefficients (MFCCs) [31-32]for analysis. It successfully predicted seizures up to an hour before they occurred, achieving high accuracy rates of 95.38% on the CHB-MIT-EEG dataset and 96.05% on the Siena-EEG dataset. Furthermore, the researchers tested a Graph Synthesizing Network (GSN) on the same datasets which created EEG graphs tailored to each subject. This method aimed to customize the seizure prediction model [33-34] to better match the unique biological patterns of individual patients, improving the model's accuracy and reducing incorrect predictions. These personalized graphs were key in providing customized predictions that adapted to each patient's specific brain activity.

III. METHODOLOGY

The research focuses on developing a robust model for seizure prediction in temporal lobe epilepsy using memory-based deep learning techniques and continual learning strategies. The approach involves several stages, from data preprocessing to model training and evaluation, incorporating mathematical formulations that are fundamental to the methods used.

A. DATA ACQUISITION AND PREPROCESSING

The dataset used in this study is sourced from the UCI repository, containing EEG signals measured in microvolts. There is total 11500 instances and 179 columns including the target variable which represents the class of epilepsy. The signals were recorded over a time period and are characterized by noise due to body movements and environmental factors. To prepare the data for modeling, several preprocessing steps were undertaken, including noise removal using hybrid filters. Mathematically, the effectiveness of noise removal is evaluated using the Signal-to-Noise Ratio (SNR), which is calculated as:

$$SNR = 10 * log_{10}(Signal\ Power/Noise\ Power) \qquad (1)$$

Here, Signal Power and Noise Power are calculated in volts. The SNR provides a quantitative measure of the filter's ability to preserve the signal while minimizing noise. Among the filters tested, the Butterworth-Wavelet denoising filter achieved the highest SNR of 24.11 dB, indicating superior performance. This filter combines a Butterworth filter for high-frequency noise removal with a wavelet transform for low-frequency noise reduction. The combination is mathematically expressed as:

$$Y(t) = W^{-1}(F\ (W(X(t)).H(f))) \qquad (2)$$

Where, $Y(t)$ is denoised EEG signal in time domain, $X(t)$ is original EEG signal in time domain, W^{-1} is inverse wavelet transform operator, F is Fourier transform operator, $H(f)$ is transfer function of the Butterworth filter.

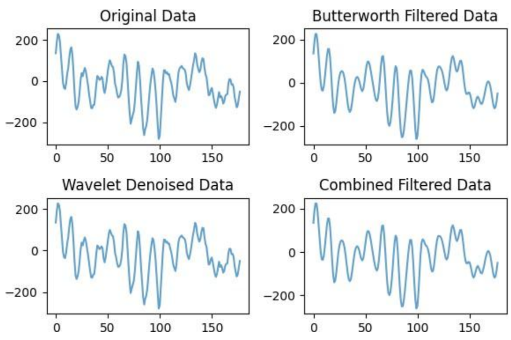

FIGURE 1. **Original Signal vs Filtered Signal.**

Figure 1 above represents the filtered signal of each phase from the best performing hybrid filter which is Butterworth Wavelet denoising in our case achieving the SNR value of 24.11 dB.

B. MODEL TRAINING

The denoised EEG signals were used to train various deep learning models, including Recurrent Neural Networks (RNNs) and transfer learning models such as DenseNet, GoogleNet etc. These models are chosen (as shown in Table I) for their ability to capture temporal dependencies in sequential data, which is crucial for accurate seizure prediction.

The training process involves minimizing a loss function that quantifies the difference between the predicted and actual outputs. For each model, we optimized the parameters using backpropagation and gradient descent algorithms. The loss function for the models can be expressed as:

$$Loss = \frac{1}{N}\sum(y_i - \hat{y}_i)^2 \qquad (3)$$

Where, y_i is the actual output, $\hat{y}_i$ is the predicted output and N is the number of training samples.

TABLE I
HYPERPARAMETERS

Parameter	Description
Training/Test Split	80% Training and 20% Testing
Validation	10% of the training
Batch	64
Epochs	100
Optimizer	Adam
Learning Rate	0.001
Adam Parametrs	B1=0.9;B2=0.999
Loss Function	Mean Squared Error (MSE)
Activation Functions	ReLU and Sigmoid
Dropout Rate	0.3
Transfer Learning Models	DenseNet, GoogleNet
Pretraining Dataset	ImageNet

C. CONTINUAL LEARNING WITH ELASTIC WEIGHT CONSOLIDATION (EWC)

To address the challenge of catastrophic forgetting in our deep learning models, Elastic Weight Consolidation (EWC) was implemented. EWC helps the model retain previously learned knowledge while adapting to new information.

The Fisher Information matrix is used to calculate the importance of each parameter in the model. This is the first phase of EWC component which can also be observed in figure 3. It is computed as:

$$FI = [Gradient(loss)]^2 \qquad (4)$$

The fisher information value as mentioned above in equation (4) is crucial for estimating the regularization term in the EWC loss function. The EWC loss function, which is the second phase of continual learning component of our model adds a regularization term to the standard loss function to penalize changes to important parameters, as defined by:

$$EWC\ term = \frac{1}{2}\sum i\ (Fisher\ Information_i * Parameter_i - Previous\ Parameter_i)^2 \qquad (5)$$

The EWC loss function becomes:

$$EWC\ loss = \lambda_{EWC} * \text{EWC Term} \qquad (6)$$

Where, λ_{EWC} is a hyperparameter controlling the strength of the regularization. This function ensures that our model can learn new tasks without significantly deviating from the previously learned ones.

D. PROPOSED ARCHITECTURE AND DESIGN

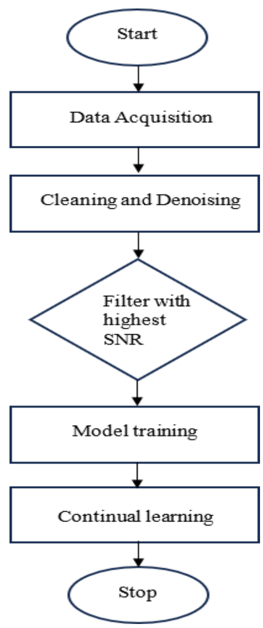

FIGURE 2. **Workflow Model**

Figure 2 shows the basic flow for implementing required models. As it can be witnessed from the figure that the implementation undergoes four basic phases namely data acquisition, preprocessing, training and continual learning.

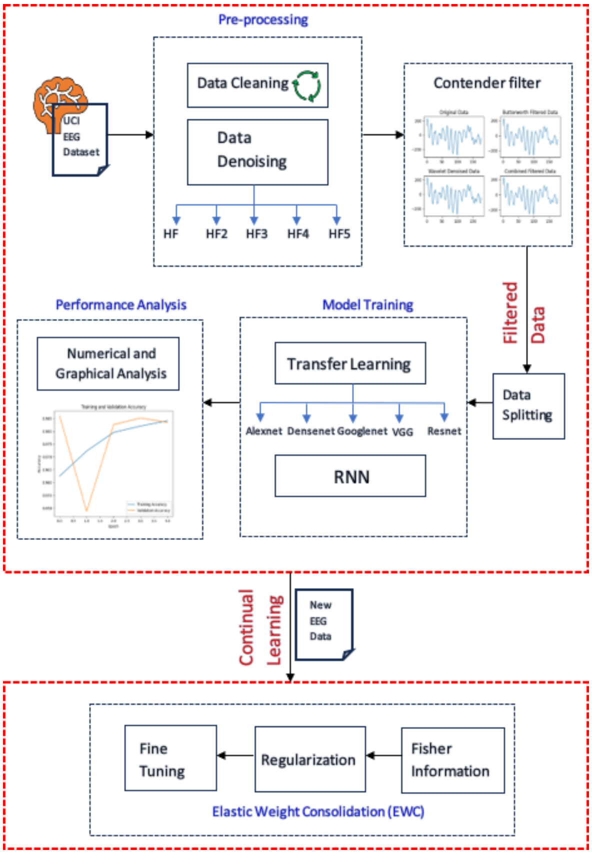

FIGURE 3. **Model Design**

The proposed model integrates both deep learning and continual learning techniques to enhance seizure prediction performance. The model design is divided into two main components: memory-based learning using transfer learning and RNN, and continual learning using EWC as can be seen from figure 3 above.

E. ALGORITHM

1. Import necessary packages and libraries
2. Import dataset
3. Data De-noising using filters
4. X-> input data

   y-> target value
5. Setup architecture Model = initialize_model ()
6. for epoch in range (num_epoch):

execute (parameters)

training (feature_parameter, target_parameter)

7. Elastic Weight Consolidation

   fisher information (model, X, y)

   EWC_loss (model, fisher information, previous parameter, lambda)

   Fine_tune ()
8. Compile and train again
9. Statistical and performance metrics analysis
10. Graphical analysis

F. PERFORMANCE EVALUATION

The model's performance is evaluated using various metrics such as accuracy, precision, recall, and F1-score. These metrics were used in both the components of architecture i.e., model training and continual learning.

1) ACCURACY

Equation (6) represents the overall accuracy of a model which is the proportion of prediction that the model got right. It should be as high as possible.

$$Accuracy = \frac{TN+TP}{TN+FN+FP+TP} \tag{7}$$

2) RECALL

Equation (7) represents the Recall or TPR which is the proportion of positive cases that the model identified correctly. It should also be as high as possible to validate the accuracy of working model.

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

3) PRECISION

Equation (8) shows the precision which is the proportion of predicted positive cases where the true label is actually positive. Similar to accuracy and recall it should also be high to indicate the performance of good model.

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

where, TP, TN, FP and FN are True Positive, True Negative, False Positive and False Negative respectively.

4) F-1 SCORE

Equation (9) gives the F-1 score which is derived from precision and recall together. It balances the trade-off between precision and recall. Its value varies from 0 to 1. The higher the value, higher will be the performance.

$$F1\ Score = \frac{2*Precision*Recall}{Precision+Recall} \tag{10}$$

IV. RESULTS AND DISCUSSION

A. NUMERICAL ANALYSIS

The numerical analysis will explore all the insights related to numbers and statistics. This will help in understanding filter accuracy like “Signal to noise ratio” and performance metrics like accuracy.

1) SIGNAL TO NOISE RATIO (SNR)

The Signal-to-Noise Ratio (SNR) is an important measure in signal processing that shows how strong the main signal is compared to unwanted noise. If the SNR is high, it means the signal is strong compared to the noise but if it is low, the noise is more significant. SNR is really important when we are talking about filters because it affects how well a filter can remove noise from a signal while keeping the important parts of the signal intact.

Signals often get mixed up with noise, which can mess up the quality of the signal. This can make it hard to understand or pick out the important parts of the signal. Filters are like tools we use to clean up signals by getting rid of unwanted noise while keeping the parts we care about. When designing filters, we try to make sure they do a good job of keeping the SNR high which means the filtered signal still looks a lot like the original one but without all the noise getting in the way. The expression for calculating SNR is already given in equation above.

Based on the range of SNR value achieved via filter, it decides how well it removes the noise from signal that might generate undesirable result if not pre-processed for model training. While dealing with time series signal it should always be de-noised using suitable filter to further process in a better manner

TABLE II
SNR TABULATION OF HYBRID FILTER

S.No.	Hybrid Filter	SNR (in dB)
1.	Gaussian-Butterworth	7.06
2.	Chebyshev Wavelet denoising	21.22
3.	Chebyshev-Bessel	22.37
4.	Daubechies-Wiener	7.73
5.	Butterworth Wavelet denoising	24.11

Table 2 above shows the SNR values obtained through different hybrid filters. There are four ranges of SNR value which dictates different significance. The SNR value greater than 20 dB shows excellent signal quality, thus resulting in clear and accurate data with minimal distortion and interference. As it can be seen in table 1, three hybrid filter managed to achieve SNR value more than 20 dB among which Butterworth wavelet denoising has the highest SNR. The second SNR range varies between 10 dB to 20 dB which shows good signal quality although there may be some noticeable background noise and interference. The third SNR range varies between 0 dB to 10 dB which indicates fair to poor signal quality. Hence in this type, the signal is somewhat weak compared to the noise leading to noticeable degradation in data. From the table we could observe that there are two such filters that has the same SNR value. The fourth SNR range indicates very poor signal quality which are also negative SNR and hence should never be considered for processing it for any desired task.

2) PERFORMANCE METRICS

The memory-based algorithms used in our work are all classifiers and hence metrics like accuracy, precision, recall and f-1 score were used to validate the performance of the model.

TABLE III
PERFORMANCE ANALYSIS

S.No.	Model	Accuracy (in %)	F-1 score	Computational Cost	Inference Time (ms)
1.	Alexnet	94.1	0.94	Low	22
2.	Densenet	96.7	0.91	High	25
3.	Googlenet	97.5	0.97	Medium	18
4.	VGG	97.0	0.97	High	35
5.	Resnet	97.5	0.97	Medium	20
6.	RNN	87.8	0.70	Medium	15

The table 3 above tabulates the accuracy and f-1 score of each implemented model. It can be observed from the table that among the memory-based learning i.e., transfer learning and RNN, transfer learning performs better than RNN. Also, all the algorithms which all are suitable for epilepsy prediction has achieved very high accuracy so it all depends on the other factors to choose the one which is new and could be reliable for continual learning. Also, Statistical analysis based on 10-fold cross-validation showed that GoogLeNet achieved an accuracy of 97.46% ± 0.11 (95% CI). A paired t-test comparing GoogLeNet and ResNet yielded $p < 0.0001$, confirming the performance difference is statistically significant.

Alexnet being the old pre-trained model is not feasible for considering. From the remaining model, Googlenet and VGG has performed almost the same. Googlenet is really good at predicting epilepsy because it is great at finding important features in EEG signals. It uses special modules called inception modules to capture different patterns in the signals, which helps it work well with different kinds of epilepsy. In contrast, VGG isn't as good at finding all the important features because it mostly uses big filters. The Inception module is a key architectural component introduced in GoogleNet (Inception v1) to improve the efficiency and performance of deep convolutional neural networks (CNNs). It enables the network to capture multi-scale spatial information in a computationally efficient way.

Also, Googlenet is smart about using its parameters which helps prevent it from getting too focused on the data it has been trained on so it can predict well even with limited EEG data. It doesn't need a lot of memory so it can work on devices with limited resources like wearable EEG monitors and still make predictions quickly. Another good thing about Googlenet is that it can understand epilepsy patterns even when they look a bit different which makes it really accurate at predicting seizures. Overall, Googlenet is a great choice for predicting epilepsy because it's smart, efficient and works in real-time. Based on these positive features of Googlenet over VGG we considered Googlenet for further processing it using continual learning.

B. GRAPHICAL ANALYSIS

The graphical analysis will explore all the insights visually and pictorially. This will help in understanding the performance of different segments of the work carried out to predict epilepsy.

1) FILTER ANALYSIS

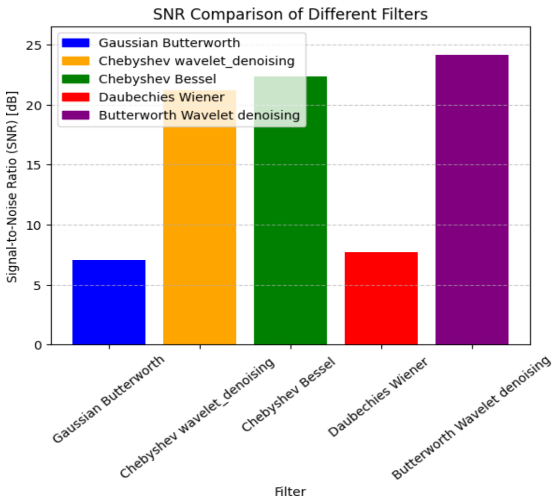

FIGURE 4. Hybrid Filter SNR Comparison

Figure 4 above shows the comparison analysis of different hybrid filter carried out in pre-processing phase. The output of the best working filter is given below for first five signal which consist of both epileptic and non-epileptic situation.

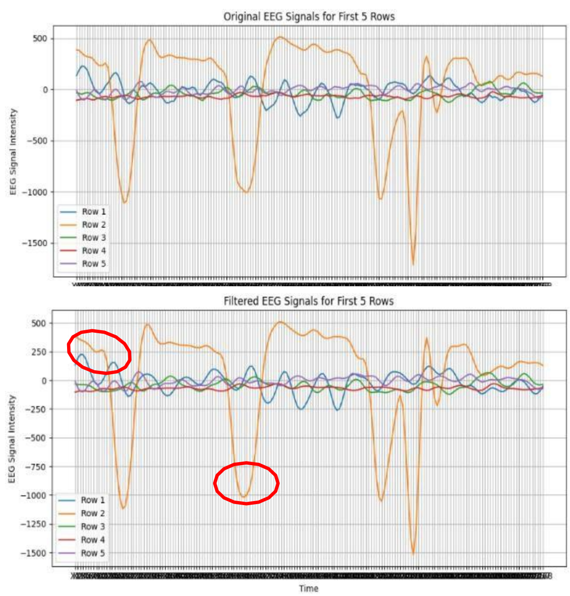

FIGURE 5. Butterworth Wavelet denoising signal Analysis

From figure 5, we could observe the filtered signal out of Butterworth wavelet denoising filter in the pre-processing stage. The glitch and sudden fluctuations were smoothed which has been highlighted as seen in the figure. This shows the removal of high frequency noise from the EEG signal.

2) PERFORMANCE ANALYSIS

Figure 6 demonstrates the performance of all the models implemented. By analyzing all the metrics, we could observe that Googlenet performs very well in all aspects. The other analysis of the best working model is given below to further validate its supremacy.

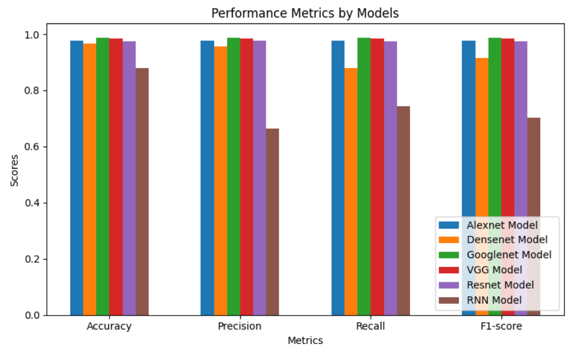

FIGURE 6. Metric Analysis

Figure 7 represents the validation curve for the Googlenet model. The steep drop in training loss indicates that the model learns a significant amount about the data in the first epoch. This is be due to the model being well-suited to the task. Having a validation loss that is consistently low and parallel to the x-axis suggests that the model is not overfitting and is performing consistently across both the training and validation datasets. This is a good sign in medical applications such as epilepsy prediction where the reliability of the model is crucial. The closeness of the training and validation loss values especially after the first epoch indicates that the model is generalizing well. This means it is likely to perform similarly on unseen data which is critical for a model used in medical diagnostics.

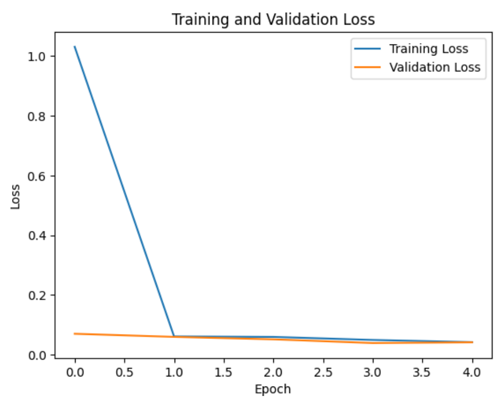

FIGURE 7. Metric Analysis

Figure 8 represents ROC curve for the Googlenet model. The AUC is 1 which is an ideal score indicating that the model can perfectly discriminate between positive and negative cases. The shape of the ROC curve indicates that for nearly all thresholds the model has a high true positive rate and a low false positive rate.

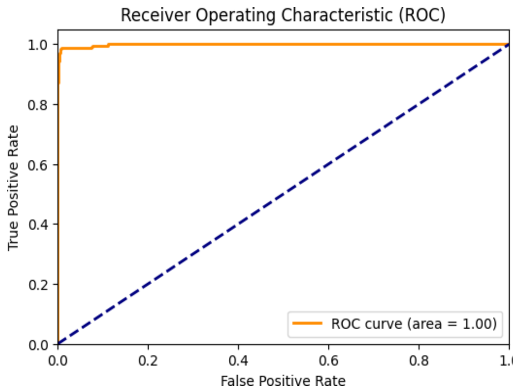

FIGURE 8. **Metric Analysis**

3) CONTINUAL LEARNING ANALYSIS

Figure 9 represents the training and validation accuracy before and after continual learning. The rapid increase in training accuracy before EWC suggests that the model is learning the training data quickly. However, this might also indicate overfitting, as the validation accuracy does not increase as sharply. The more gradual increase in training accuracy with EWC indicates that the model is learning more cautiously which can prevent overfitting and lead to better generalization. The validation accuracy before EWC shows more fluctuations, especially around epoch 2, indicating potential instability or overfitting. The validation accuracy with EWC shows a more stable trend, suggesting better generalization to unseen data. The training accuracy is consistently higher than the validation accuracy, suggesting some overfitting. The training and validation accuracies are closer, indicating that the model is generalizing better to the validation set.

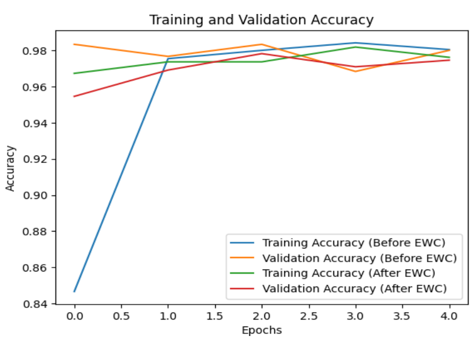

FIGURE 9. **Training and Validation Accuracy**

Figure 10 shows the validation loss of a model before and after applying continual learning. The initial higher loss with EWC might indicate that the regularization imposed by EWC initially slows down the learning process. The smoother decline in the validation loss after EWC suggests that the model might be learning more consistently without abrupt changes, potentially indicating better stability. The model without EWC shows a faster reduction in validation loss initially but experiences a slight increase in loss at epoch 3 which could suggest overfitting or instability at that point. The model with EWC does not show such an abrupt increase indicating more stable learning. Although the final validation loss is slightly lower for the model without EWC, the model with EWC exhibits more consistent learning, which might be beneficial for continual learning scenarios where stability and prevention of catastrophic forgetting are crucial.

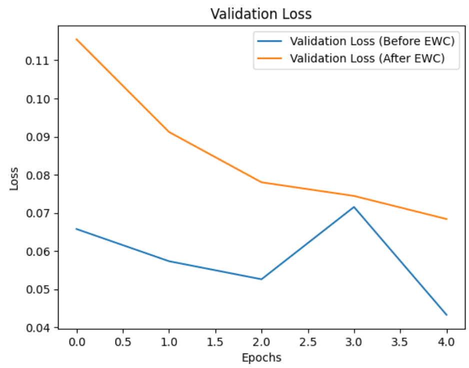

FIGURE 10. **Validation Loss for EWC**

Figure 11 indicates variation of EWC loss and accuracy with respect to different values of lambda to show the impact of regularization. The increasing EWC loss indicates that the model is accumulating a significant amount of penalty over time for deviating from the previously learned weights. This suggests that the model is making adjustments to its parameters that the EWC mechanism penalizes to maintain performance on previous tasks. The constant accuracy suggests that the model's predictive performance on the current task remains unchanged, despite the increasing EWC loss. This indicates that the model is maintaining its performance but is being heavily regularized by EWC. The substantial increase in EWC loss might suggest that the regularization parameter (λ) for EWC could be too high, leading to over-regularization. This prevents the model from adequately adapting to new data, thereby keeping the accuracy constant.

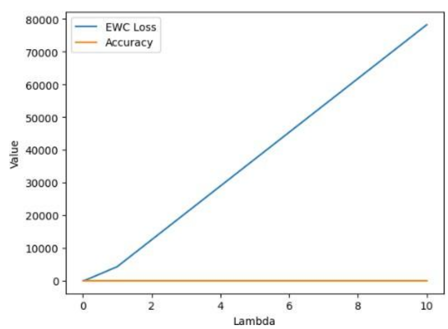

FIGURE 11. **Regularization under EWC**

4) COMPARATIVE ANALYSIS WITH EXISTING MODELS

Table 4 shows the performance analysis of existing model and the proposed model based on the accuracy on UCI EEG dataset. As it can be inferred from the above, Googlenet along with the Custom Elastic Weight Consolidation achieves the highest accuracy of 98.7% which is higher when compared to other. Not only this, the continual learning technique used would further help us to obtain a stable performance while training the new dataset.

TABLE IV
PERFORMANCE ANALYSIS OF EXISTING AND PROPOSED MODEL

S.No.	Model Used	Accuracy (in %)
1.	CNN, LSTM	93.9
2.	CNN, MLP	-
3.	CNN-LSTM Hybrid	97.5
4.	Googlenet-Continual Learning **(Proposed)**	98.7

GoogLeNet provides a good compromise between efficiency and accuracy. Because of its Inception modules, which enable parallel, multi-scale feature extraction without unduly increasing depth, it has a substantially smaller model size (~6.8M parameters) and lower computational overhead (~1.5 GFLOPs) even though it achieves performance comparable to ResNet and VGG. Because of this, it is especially well-suited for real-time seizure detection and deployment on edge devices, such as mobile health platforms or portable EEG monitors.On the other hand, because VGG and ResNet require more memory and power and are computationally heavy, they are less suitable for low-resource environments.Also, despite being smaller, DenseNet features a more intricate connectivity pattern, which lengthens training times and increases memory access. Although RNNs are good at handling sequential input, their computational cost during training and inference prevents them from being used in real-time without specific optimization.

V. CONCLUSION

Creating a predictive model for epilepsy using deep learning and continual learning is a big step forward in epilepsy management. During this research work, we have looked at how deep learning like transfer learning and recurrent neural networks (RNNs) can understand the complex patterns in epileptic seizures. By using deep learning, we have been able to find important information in EEG signals and make better predictions about seizures. Among RNN and transfer learning, transfer learning proved to be more accurate model for classifying seizure in temporal lobe epilepsy. By using Googlenet as the transfer learning model we achieved the highest accuracy of 97% in training our dataset. Also, by using Elastic Weight Consolidation (EWC) as the continual learning technique with a learning rate scheduler our model enhanced the accuracy to 98.7%. Unlike older models that stay the same our continual learning model can learn from new information while still remembering what it has learnt before. This flexibility has made our model better at predicting seizures and giving timely alerts and help to patients and caregivers.Current seizure prediction models have a number of drawbacks. They have trouble identifying long-term dependencies in EEG signals because standard models, such as RNNs and LSTMs, have trouble understanding intricate temporal patterns. Furthermore, these models frequently don't adjust to changing patient data, including modifications in medicine or the course of a disease. Many models still have trouble with noisy EEG data from artifacts, even though methods like Elastic Weight Consolidation (EWC) can aid in continuous learning and allow models to adjust without losing prior information. Additionally, models like as CNNs and ResNets are not feasible for real-time applications on edge devices due to their high processing requirements. Lastly, existing models often face challenges in generalizing across different patient populations due to the variability of seizure patterns, though transfer learning and domain adaptation techniques are being explored to address this. Overall, our work shows how advanced machine learning techniques can improve how we care for people with epilepsy, giving them better support and management options.

Compliance with Ethical Standards

No study related to human or animal was involving. Standard data in the public repository was used for kinds of results.

Competing Interest

The authors declare that they have no conflict of interest.

REFERENCES

[1] B Jaishankar, Ashwini AM, Vidyabharathi D and L Raja, “A novel epilepsy seizure prediction model using deep learning and classification”, Elsevier, Vol. 4, 2023.

[2] AS Zandi, GA Dumont, M Javidan and R Tafreshi, “An Entropy-Based Approach to Predict Seizure in Temporal Lobe Epilepsy Using Scalp EEG”, IEEE Engineering in Medicine and Biology Society, 2009.

[3] Z Zhang and KK Parhi, “Seizure Prediction using Polynomial SVM Classification”, 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5748-5751, 2015.

[4] SM Usman, S Khalid and Z Bashir, “Epileptic Seizure Prediction using Scalp Electroencephalogram Signals”, Elsevier, Vol. 41(1), 2021.

[5] MS Farooq, A Zulfiqar and S Riaz, “Epileptic Seizure Detection Using Machine Learning: Taxonomy, Opportunities and Challenges”, PubMed, DOI: 10.3390/diagnostics13061058, 2023.

[6] ME Saab and J Gotman, “A System to Detect the Onset of Epileptic Seizures in Scalp EEG”, Elsevier, Vol. 116(427-442), 2004.

[7] N Wang and MR Lyu, “Extracting and Selecting Distinctive EEG Features for Efficient Epileptic Seizure Prediction”, IEEE, Vol. 19(1648-1659), 2014.

[8] R Chen and KK Parhi, “Seizure Prediction using Convolutional Neural Networks and Sequence Transformer Networks”, IEEE, 2021.

[9] A Shoeibi et.al, “An Overview of Deep Learning Techniques for Epileptic Seizures Detection and Prediction based on Neuroimaging Modalities: Methods, Challenges and Future Works”, Computers in biology and medicine, 2022.

[10] A Shoeibi et. Al, “Epileptic Seizure Detection Using Deep Learning Techniques:: A review”, PubMed, 2021.
[11] A Abdelhameed and M Bayoumi, “A Deep Learning Approach for Automatic Seizure Detection in Children with Epilepsy”, Frontier Computing Neuroscience, Vol. 15, 2021.
[12] S Hou, X Pan, CC Loy, Z Wang and D Lin, “Learning a Unified Classifier Incrementally via Rebalancing”, CVF Conference on Computer Vision and Pattern Recognition, pp. 831-839, 2019.
[13] V Kiri, “An Explainable Multi-Modal Neural Network Architecture for Predicting Epilepsy Comorbidities Based on Administrative Claims Data”, Academia, 2020.
[14] ZG Yap, MR Popovic and SS Khan, “Supervised and Unsupervised Deep Learning Approaches for EEG Seizure Prediction”, arXiv, DOI: https://doi.org/10.48550/arXiv.2304.14922, 2023.
[15] K Rasheed, J Qadir, TJ O’Brien, L Kuhlmann and A Razi, “Generative Model to Synthesize EEG Data for Epileptic Seizure Prediction”, arXiv, 2020.
[16] M Hartmann, K Schindler, TA Gebbink, G Gritsch and T Kluge, “PureEEG: Automatic EEG artifact removal for epilepsy monitoring”, Elsevier, 2014, pp. 479-490 (44).
[17] M Hartmann et. al, “Seizure detection with deep neural network for review of two-channel electroencephalogram”, Epilepsia, 2022, doi: 10.1111/epi.17259.
[18] LS Remvig et. al, “Detecting temporal lobe seizures in ultra long subcutaneous EEG using algorithm-based data reduction”, Elsevier, 2022, pp. 86-93 (142).
[19] F Furbass et. al, “Automatic multimodal detection for long-term seizure documentation in epilepsy”, Elsevier, 2017, pp. 1466-1472 (128).
[20] I Bhattacherjee, “Epileptic seizure prediction using machine learning techniques on real-time EEG signals”, IEEE, 2021.
[21] M Masum, H Shahriar and HM Haddad, “Epileptic seizure detection for imbalanced datasets using an integrated machine learning approach”, IEEE, 2020.
[22] S Sharma and A Arora, “Detection of epileptic seizure using machine learning”, IEEE, 2022.
[23] C Salvador et. al, “Epileptic seizure prediction using EEG peripheral channels”, IEEE, 2023.
[24] S Kumar et. al, “Machine Learning for seizure prediction: a revamped approach”, IEEE, 2015.
[25] TI Rohan, MS Uddin, YM Islam and S Roy, “Efficient approach to detect epileptic seizure using machine learning models for modern healthcare system”, IEEE, 2020.
[26] B Seifi, M Barfi and M Esmaeilpour, “ECG-based prediction of epileptic seizure using machine learning model”, IEEE, 2022.
[27] Y Qin et. al, “Patient specific seizure prediction with scalp EEG using convolutional neural network and extreme learning machine”, IEEE, 2020.
[28] X Wang, C Zhang, T Karkkainen, Z Chang and F Cong, “Channel increment strategy based 1D convolutional neural network for seizure prediction using intracranial EEG”, vol. 31, IEEE, 2023.
[29] C Chen et. al, “Epileptic seizure prediction based on EEG by auto-machine learning”, IEEE, 2022.
[30] T Dissanayake, T Fernando and S Sridharan, “Geometric Deep Learning for subject Independent Epileptic Seizure Prediction Using Scalp EEG Signals”, vol. 26 (2), IEEE, 2022.
[31] Q Huo, X Luo, ZC Xu and XY Yang, “Machine Learning applied to Epilepsy: Bibliometric and Visual Analysis from 2004 to 2023”, Frontiers Neurology, Vol. 15, 2024.
[32] BR Nayana, MN Pavithra, S Chaitra, TN Bhuvana Mohini, T Stephan, V Mohan, N Agarwal, “EEG-based neurodegenerative disease diagnosis: comparative analysis of conventional methods and deep learning models”, Scientific Reports, 5(1), pp.1-3, 2025.
[33] N Delfan, M Shahsavari, S Hussain, R Damaševičius, UR Acharya, “A Hybrid Deep Spatiotemporal Attention-Based Model for Parkinson’s Disease Diagnosis Using Resting State EEG Signals”, International Journal of Imaging Systems and Technology, 34(4), e23120, 2024.
[34] H Lin, J Fang, J Zhang, X Zhang, W Piao, Y Liu, “Resting-state electroencephalogram depression diagnosis based on traditional machine learning and deep learning: A comparative analysis”, Sensors, 23; 24(21): 6815, 2024.

BIKAS CHANDRA SAHOO currently pursuing the M.Tech. degree in Computer Science engineering with specializations in Big Data Analytics with the Vellore Institute of Technology, Chennai, India. His research interests include Artificial Intelligence, Machine Learning, data analytics and deep learning and its applications.

SHRUTI MISHRA (Member, IEEE) received the Ph.D. degree in computer science and engineering from Siksha ’O’ Anusandhan University, Bhubaneswar, Odisha, India. She had been an Associate Professor with the Department of Computer Science and Engineering, Vignana Bharathi Institute of Technology, Hyderabad. She is currently an Assistant Professor (Senior) with the Centre of Advanced Data Science, Vellore Institute of Technology, Chennai. She has more than 30 papers in both national and international to her credit along with three books. She has guided more than 40 postgraduate and undergraduate students. She is the guest editor of many reputed publishers, such as Elsevier. She has also served as a reviewer for many reputed journals and conferences.

SANDEEP KUMAR SATAPATHY (Member, IEEE) received his Post-Doctoral Fellow from Yonsei University, Seoul, South Korea. He received his Ph.D. degree in data mining and machine learning. His Ph.D. thesis include a detailed classification of brain EEG signals using machine learning techniques. He was an Associate Professor with the Department of Computer Science and Engineering and the Head of the Department of Information Technology, Vignana Bharathi Institute of Technology, Hyderabad. Currently, working as an Associate Professor at Centre for Advanced Data Science, VIT University, Chennai. He is highly engrossed into the areas of deep learning, image processing, and machine learning. He has many research publications to his credit, such as more than 40 research articles, three books, and many book chapters in various peer-reviewed journals. He has guided more than 15 master’s thesis. He has also authored two books, such as Frequent Pattern Discovery from Gene Expression Data: An Experimental Approach (Elsevier) and EEG Brain Signal Classification for Epileptic Seizure Disorder Detection (Elsevier). He has been a member of various academic committees within the institution. He is a member of many professional organizations and society. He has been an active reviewer of various peer-reviewed journals and prestigious conferences. He has also reviewed many research articles and books in Elsevier for possible publication.

SACHIN KUMAR (Senior Member, IEEE) currently working as an Associate Professor, Akian College of Science and Engineering, American University of Armenia. Dr. Kumar earned his Ph.D. in machine learning from the Indian Institute of Technology in Roorkee, India. He completed his master’s and bachelor’s degrees in computer science and engineering from the State Universities of India. Prior to joining AUA, Dr. Kumar was a leading scientist in the Laboratory of Big Data and Machine Learning at South Ural State University (SUSU, National Research University) in Chelyabinsk, Russia, where he also served as a project leader for and participant in various research projects funded by the Russian Science Foundation from March 2019- June 2024. In December 2022, he received the SUSU Recognition Award in the “Science” category. From 2022 to 2023, he was also a visiting professor at the College of Information Business System at the National University of Science and Technology MISiS in Moscow, Russia, where he previously worked as an assistant professor and post-doctoral researcher from December 2017 to February 2019. He is also a member of the Scientific Dissertation Council in the field of information technology in the Chelyabinsk region of Russia.

Dr. Kumar’s research interests include smart healthcare, Internet of things, intelligent transportation systems, and natural language processing. He enjoys teaching a variety of core and elective courses related to computer science and data science.