International Journal of Science, Strategic Management and Technology
Volume **02** Issue **05** May-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Multi-Class EEG-Based Epileptic Seizure Classification using Hybrid Deep Learning Architectures

Author Details:

Sanjivani Adusl, Aditya Jain, Abhijeet Kolhe, Akshat Patil,Arya Manve

[1] Artificial Intelligence and Data Science /Vishwakarma Institute of Technology / Pune, India
[2] Artificial Intelligence and Data Science /Vishwakarma Institute of Technology / Pune, India
[3] Artificial Intelligence and Data Science /Vishwakarma Institute of Technology / Pune, India

4 Artificial Intelligence and Data Science /Vishwakarma Institute of Technology / Pune, India

5 Artificial Intelligence and Data Science /Vishwakarma Institute of Technology / Pune, India

Corresponding Author Email: jain.aditya23@vit.edu

https://doi.org/10.55041/ijsmt.v2i5.192
Cite this Article: Adusl, S., Jain, A., Kolhe, A., Patil, A. & Manve, A. (2026). Multi-Class EEG-Based Epileptic Seizure Classification using Hybrid Deep Learning Architectures. International Journal of Science, Strategic Management and Technology, *02*(05).
https://doi.org/10.55041/ijsmt.v2i5.192

Abstract—

Epilepsy affects roughly 50 million people worldwide, yet reliable automated seizure detection remains an open problem. Manual EEG interpretation is slow, subjective, and requires specialist expertise that is scarce in many clinical settings. We built an automated five-class seizure classification system and evaluated three deep learning architectures—a 1D Convolutional Neural Network (1D-CNN), a Bidirectional LSTM (BiLSTM), and a CNN-LSTM hybrid—against traditional baselines (Random Forest, SVM, Decision Tree) on the Bonn University EEG dataset. The 1D-CNN achieved the highest accuracy at 95.5% (precision 95.6%, recall 95.5%, F1 95.5%), outperforming all other models. A Streamlit web application was developed alongside the models to support real-time EEG upload and seizure prediction. Our results suggest that convolutional feature extraction is particularly effective for this task, even when compared to architectures designed to model temporal sequences.

Keywords— EEG Signal Processing; Epileptic Seizure Detection; Deep Learning;1D Convolutional Neural Network; Bidirectional LSTM, CNN–LSTM Hybrid, Bonn EEG Dataset

I. INTRODUCTION

Epilepsy is a chronic neurological disorder marked by recurrent, unprovoked seizures caused by abnormal electrical activity in the brain. Over 50 million people live with epilepsy globally, making it one of the most burdensome neurological conditions in terms of patient quality of life, caregiver load, and healthcare costs [1]. Despite available treatments, roughly 30% of patients do not respond adequately to medication, which makes early and accurate seizure detection a practical priority rather than just an academic one.

EEG is the primary diagnostic tool for epilepsy, offering high temporal resolution of electrical brain activity across scalp electrodes. The problem is that reading EEG is hard. Neurologists spend hours visually scanning multichannel waveforms to identify ictal and interictal patterns—work that is inherently time-consuming and varies between reviewers. In settings where trained epileptologists are not available, this bottleneck is even more consequential.

Prior automated approaches have drawn on a wide range of signal processing and machine learning methods. Classical classifiers like SVMs, Random Forests, and Decision Trees show moderate performance, but they depend heavily on hand-engineered features and tend to struggle when applied across different patient populations or recording setups. This has pushed the field toward deep learning, where models can learn feature representations directly from the signal.

This work focuses on five-class epileptic seizure classification using the Bonn University EEG dataset, which includes five distinct signal types (Sets A–E): healthy subjects with eyes open, healthy subjects with eyes closed, seizure-free interictal activity from the epileptogenic zone, seizure-free interictal activity from the hippocampal formation, and full ictal activity. Distinguishing all five classes is considerably harder than binary detection, but it also gives more useful clinical information.

Our objective was to build a robust five-class EEG classification system combining time-frequency feature extraction (DWT and PSD) with three deep learning architectures: a 1D-CNN for local temporal pattern learning, a BiLSTM for capturing sequential dependencies in both directions, and a CNN-LSTM hybrid that combines both approaches.

The main contributions are: a full five-class evaluation (rather than the simplified binary setups common in prior work), a unified experimental comparison of six models spanning both classical and deep learning approaches, a 1D-CNN achieving 95.5% accuracy and F1-score, and integration of the best model into a Streamlit web application that accepts raw EEG input and returns predictions with confidence scores.

II. LITERATURE REVIEW

Automated EEG seizure detection has been studied for decades. The field has broadly followed the same arc as machine learning more generally, starting with hand engineered features fed into classical classifiers, moving through shallow neural networks, and more recently shifting toward end-to-end deep learning. What follows is an overview of the key directions in this space and where existing work still falls short.

Traditional Machine Learning Approaches

Early systems extracted handcrafted features from EEG's time domain descriptors like mean, variance, skewness, and zero-crossing rate and fed them into standard classifiers. Frequency domain features derived from FFT and PSD analysis were also widely used, decomposing the signal into clinically recognized sub bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (>30 Hz). These sub band features tend to correlate with distinct brain states and showed reasonable discriminative power in early experiments.

Andrzejak et al. [2] whose dataset forms the basis of this study showed early on that non linear signal descriptors could differentiate ictal from interictal and healthy EEG. Subasi [3] and Übeyli [4] followed with wavelet based features combined with Mixture of Expert models and Multilayer Perceptrons, reporting binary classification accuracies between 94 to 98% on the same dataset. Shoeb and Guttag [5] took a patient-specific approach using SVMs with time and frequency features and achieved high sensitivity in a clinical setting.

Random Forest has also seen wide use in this space. Acharya et al. [6] showed that combining higher order spectral and non linear features with ensemble methods improves generalization. In our experiments, Random Forest reached 85.0% accuracy and 84.9% F1, a solid, but clearly limited by the feature engineering ceiling that all classical models hit when faced with the complex, non stationary nature of ictal EEG.

SVMs have probably been the most studied classical method in this domain. They perform well in binary settings but struggle with multi class problems of kernel choice and class imbalance sensitivity both become harder to manage. In our five class evaluation, SVM reached only 76.0% accuracy with a 73.1% F1 score, highlighting the gap between binary optimized classical models and the more realistic multi class setting.

The Shift Toward Deep Learning

The limitations of hand crafted feature pipelines have catalyzed a substantial shift in the EEG seizure detection literature toward deep learning architectures, which offer the compelling advantage of automated and data driven feature extraction. Convolutional Neural Networks (CNNs), originally developed for image recognition tasks, have been successfully adapted for one dimensional time series classification. Their ability to detect local temporal patterns through learned convolutional filters makes them particularly well suited to EEG signal analysis, where transient morphological features, such as spike and wave discharges carry diagnostic significance.

Acharya et al. [7] proposed a 13 layer 1D-CNN for automatic detection of focal and non focal EEG signals, reporting accuracy improvements over traditional feature-based methods. Yuan et al. [8] demonstrated that CNN architectures could extract discriminative features from raw EEG without explicit frequency band decomposition, achieving strong performance on seizure detection benchmarks. The 1D-CNN proposed in the present paper is consistent with these results, with a classification accuracy of 95.5% and F1 score of 95.5%, which is the best among all the compared methods.

Recurrent Neural Networks (RNN), in particular the Long Short-Term Memory (LSTM) based networks, have recently gained traction as a complementary deep learning approach to the analysis of EEG data, due to the sequential nature of the brain signals and the temporal dependencies between them. Tsiouris et al. [9] applied LSTM networks to long term EEG recordings for seizure prediction, demonstrating superior sensitivity compared to non recurrent baselines. The Bidirectional LSTM (BiLSTM) architecture, which processes input sequences in both temporal directions, has been shown to capture richer contextual dependencies than its unidirectional counterpart. However, as observed in the present study, the BiLSTM model achieved a classification accuracy of 79.9% and an F1 score of 79.8%, suggesting that recurrent modeling alone, without convolutional preprocessing which may be insufficient to fully exploit the spatial-temporal structure of segmented EEG data.

Hybrid architectures that combine CNN and LSTM components have attracted increasing attention as a means of leveraging the complementary strengths of both paradigms. Sainath et al. [10] demonstrated the efficacy of CNN and LSTM hybrids for sequential signal modeling in acoustic tasks, and subsequent work has extended this principle to biomedical time series. In the EEG domain, Zhang et al. [11] proposed a CNN and LSTM model for seizure detection that outperformed either architecture in isolation. The CNN-LSTM Hybrid model developed in the present work achieved a classification accuracy of 93.1% and an F1 score of 93.1%, confirming the viability of this integrated approach while revealing that, for the specific five class Bonn dataset task, the 1D-CNN alone demonstrated marginally superior aggregate performance.

Identified Gaps in Existing Literature

Quality of Previous Work and Potential Limitations Although there has been considerable advancement, there are still a few significant limitations in the prior work on automated EEG seizure detection. The most prevalent of these is the dominant attention to binary classification. It is usually formulated as seizure vs. non seizure, or ictal vs. healthy which, although of clinical relevance, does not reflect the diagnostic granularity needed for whole epileptic patient management. Most of the published deep learning approaches on the Bonn dataset consider only two or three out of the five available signal classes, thus alleviating the classification challenge and artificially enhancing the performance metrics obtained as compared to the more challenging five-class problem.

Besides, in the same experimental settings there is no clear discussion or neutral comparison of handcrafted feature based traditional approaches and deep learning based automated methods. Studies generally only recommend one paradigm or the other, which undermines the interpretability of comparative assertions. The propensity for models to be evaluated separately with no standard preprocessing routines, same train-test split, or common metrics of evaluation adds confounding noise that makes cross-study comparisons harder.

Attention to clinical deploy ability also remains limited, while numerous studies demonstrate promising offline classification accuracy, few address the practical pathway from a trained model to a usable clinical tool. The integration of trained models into interactive diagnostic interfaces capable of processing real patient EEG files and returning interpretable predictions in real time represents an underexplored dimension of applied seizure detection research.

Positioning of the Present Work

The present study directly addresses these identified gaps through three distinguishing contributions. First, the full five-class classification problem defined by the Bonn University EEG dataset encompassing Sets A through E which is tackled in its entirety, providing a more rigorous and clinically meaningful evaluation benchmark than binary or three-class reductions. Second, a unified experimental framework is employed to compare six models spanning both the handcrafted feature paradigm (Random Forest, SVM, Decision Tree with PSD and wavelet-derived features) and the automated deep learning paradigm (1D-CNN, BiLSTM, CNN-LSTM Hybrid), enabling direct and fair cross-paradigm performance assessment. Third, the highest-performing model is integrated into a deployable Streamlit web application, demonstrating a concrete step toward translating research-grade seizure detection into an accessible clinical decision support tool. Altogether, these contributions make the current work a full-fledged, practical step forward from the state-of-the-art for automated classification of EEG epileptic seizures.

Study	**Dataset**	**Classes**	**Method**	**Accuracy (%)**
Subasi (2007)	Bonn	2-class	Wavelet + MLP	94.5
Übeyli (2008)	Bonn	2-class	Eigenvector + SVM	96.2
Acharya et al. (2018)	Bonn	2-class	1D-CNN	98.0
Zhang et al. (2018)	Bonn	3-class	CNN-LSTM	95.1
Proposed Work	Bonn	5-class	CNN–LSTM Hybrid	93.1
Proposed Work	Bonn	5-class	1D-CNN	95.5

Table.2. Comparison with Previous Studies

III. METHODOLOGY

This section presents the complete methodological pipeline employed in the development of the automated EEG epileptic seizure classification system. The workflow encompasses dataset acquisition and characterization, signal preprocessing and segmentation, time-frequency feature engineering, and the design of three deep learning model architectures. Figure 1 illustrates the overall system pipeline from raw EEG input to classification output.

Dataset Description

The experiments in this study are conducted on the publicly available Bonn University EEG dataset, originally introduced by Andrzejak et al. [2], which has served as a standard benchmark for epileptic seizure classification research for over two decades. The dataset comprises five sets of EEG recordings, designated Sets A through E, each containing 100 single-channel EEG segments of 23.6 seconds duration, recorded at a sampling frequency of 173.6 Hz, yielding 4096 data samples per segment.

The five sets represent distinct neurophysiological conditions. Set A contains surface EEG recordings from five healthy volunteers with eyes open, while Set B contains recordings from the same subjects with eyes closed. Both sets were acquired using standardized electrode placement according to the international 10-20 system. Sets C and D contain intracranial depth electrode recordings from five epileptic patients during seizure-free intervals; Set C was recorded from within the hippocampal formation of the hemisphere contralateral to the epileptogenic zone, whereas Set D was recorded from within the epileptogenic zone itself. Set E contains ictal EEG segments recorded exclusively during epileptic seizures, exhibiting the characteristic high-amplitude spike-and-wave discharge morphology associated with active ictogenesis.

This five-class configuration, spanning healthy activity, seizure-free interictal states from distinct brain regions, and full ictal activity which constitutes a considerably more challenging classification problem than the binary or reduced three-class formulations commonly adopted in the literature, and provides a clinically comprehensive basis for model evaluation. The entire dataset contains 500 EEG segments in total across five balanced classes, with 100 samples in each class, providing a fair distribution that avoids the effect of class imbalance during training and testing.

Data Preprocessing

Raw EEG signals from the Bonn dataset, each comprising 4096 samples at 173.6 Hz, undergo a structured preprocessing pipeline designed to enhance signal quality, standardize input dimensionality, and improve the generalization capacity of the downstream classification models.

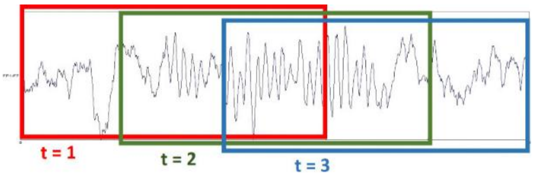


Fig. 1. EEG Segementation for Alpha Beta Theta Delta Gamma in signal

- Segmentation: Each original 4096-sample EEG recording is partitioned into non-overlapping fixed-length segments of 512 samples, yielding eight sub-segments per original file. This segmentation strategy, implemented via the “prepare_dl_data” function in the present codebase, substantially augments the effective training dataset size from 500 to 4000 samples, providing a richer and more diverse set of training examples without introducing data leakage between train and test partitions. A segment length of 512 samples was selected as it preserves sufficient temporal context to capture complete ictal morphological patterns, including spike-wave complexes and high-frequency oscillations, while remaining computationally tractable for convolutional and recurrent processing.

- Z-Score Normalization: Following segmentation, each individual EEG segment is normalized using Z-score standardization, defined as:

$$\hat{x}_i = \frac{x_i - \mu}{\sigma}$$

where $x_i$denotes the raw amplitude value at sample index i, $\mu$is the segment mean, and $\sigma$is the segment standard deviation. Z score normalization is applied on a per segment basis to eliminate inter-subject and inter-session amplitude variability, ensuring that the models learn morphological and temporal patterns rather than absolute amplitude differences. This normalization approach is particularly critical given that EEG recordings across the five sets were acquired under differing electrode configurations, scalp versus intracranial which introduce systematic amplitude offsets that would otherwise confound model training.

- Train-Test Split: The preprocessed and normalized samples are split into train and test sets with a stratified 80/20 split that maintains the class distribution in both splits. Stratified splitting is important in multi class scenarios to have an adequate representation of each class in the train and test set, otherwise the evaluation might be biased on favour of the classes in majority.

Feature Engineering

For the traditional machine learning classifiers, Random Forest, SVM, and Decision Tree discriminative features are extracted from each preprocessed EEG segment using two complementary time frequency analysis techniques: Power Spectral Density estimation via Welch's method and Discrete Wavelet Transform decomposition. These handcrafted features encode the frequency domain and multi resolution temporal characteristics of EEG signals that are known to differ systematically across the five neurophysiological states represented in the dataset. The db4 wavelet was selected after empirical evaluation due to its superior performance in capturing epileptic spike morphology compared to Haar and Symlet wavelets.

Power Spectral Density Using Welch's Method

Power Spectral Density (PSD) characterizes the distribution of signal power as a function of frequency and is among the most widely employed features in EEG analysis. In the present work, PSD is estimated using Welch's method, which improves upon the periodogram by averaging over overlapping windowed sub-segments of the signal, thereby reducing spectral variance and improving estimate reliability. Welch's method divides the input signal into Koverlapping segments of length L, applies a Hann window to each segment to reduce spectral leakage, computes the periodogram of each windowed segment, and averages the resulting periodograms:

$$\hat{S}(f) = \frac{1}{K} \sum_{k=1}^{K} |\mathrm{DFT}\{w[n] \cdot x_k[n]\}|^2$$

where $w[n]$ denotes the Hann window function and $x_k[n]$ denotes the k -th signal segment. Statistical summaries including the mean, standard deviation, maximum, minimum, and total power are computed from the resulting PSD estimate across the clinically defined EEG frequency sub-bands. These sub-band power features encode the relative contribution of delta, theta, alpha, beta, and gamma oscillations, which exhibit well-characterized differential patterns across ictal, interictal, and healthy EEG states and thereby provide strong discriminative signals for the classical classifiers.

Discrete Wavelet Transform Using the db4 Wavelet

The Discrete Wavelet Transform (DWT) provides a multi-resolution time-frequency decomposition that is particularly well-suited to non-stationary signals such as EEG, where spectral characteristics evolve dynamically over time. Unlike the Fourier transform, which provides only global frequency information, the DWT decomposes a signal into localized wavelet coefficients at multiple scales, enabling simultaneous temporal and frequency resolution.

In the present implementation, the Daubechies 4 (db4) mother wavelet is employed for DWT decomposition, a choice motivated by its established efficacy in EEG analysis due to its compact support, near-symmetry, and morphological similarity to characteristic EEG waveforms including sharp waves and spike discharges. A four level decomposition is applied to each 512 sample segment, producing one approximation coefficient sub-band and four detail coefficient sub bands. The resulting decomposition levels correspond approximately to the following frequency ranges at the 173.6 Hz sampling rate: Detail level D1 (43.4–86.8 Hz, gamma), Detail level D2 (21.7–43.4 Hz, beta-gamma), Detail level D3 (10.85–21.7 Hz, alpha-beta), Detail level D4 (5.4–10.85 Hz, alpha-theta), and Approximation level A4 (0–5.4 Hz, delta-theta).

Statistical features including mean, standard deviation, energy, entropy, and the mean absolute value of coefficients are extracted from each decomposition level, yielding a compact and informationally rich feature vector that captures both the temporal localization and the frequency composition of transient EEG events. The combined PSD and DWT feature vectors are concatenated to form the final feature representation supplied to the classical machine learning classifiers.

Model Architecture

Three deep learning architectures are designed, implemented, and evaluated for the five-class EEG seizure classification task. All models receive as input the normalized 512-sample EEG segments reshaped as one-dimensional time-series tensors of shape (512, 1), and produce a five-dimensional softmax output corresponding to the probability distribution over the five EEG classes.

1. One-Dimensional Convolutional Neural Network (1D-CNN)

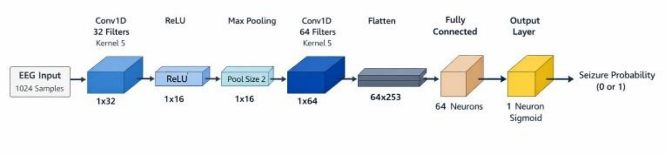


Fig. 2. CNN Workflow Architecture

The 1D-CNN architecture is designed to perform automated spatial feature extraction from raw EEG segments by learning a hierarchy of local temporal patterns through successive convolutional operations. The network architecture consists of three convolutional blocks, each comprising a one-dimensional convolutional layer, Batch Normalization, ReLU activation, and MaxPooling. The first convolutional block applies 64 filters of kernel size 3, capturing fine-grained local temporal patterns such as sharp transients and spike morphologies. The second block applies 128 filters of kernel size 3, learning higher-order combinations of the features extracted by the first block. The third block applies 256 filters of kernel size 3, encoding abstract, semantically rich representations of EEG signal structure across longer temporal extents.

Batch Normalization is applied after each convolutional layer to stabilize training dynamics, accelerate convergence, and reduce sensitivity to weight initialization. MaxPooling with a pool size of 2 follows each convolutional block, progressively reducing the temporal dimensionality of the feature maps while retaining the most salient activations. The convolutional output is subsequently flattened and passed through two fully connected layers of 256 and 128 units respectively, both employing ReLU activation and Dropout regularization with a rate of 0.5 to mitigate overfitting. The final output layer applies softmax activation over five units to produce the class probability distribution. The complete 1D-CNN model contains approximately 1.2 million trainable parameters and is trained using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss.

2. Bidirectional Long Short-Term Memory (BiLSTM)

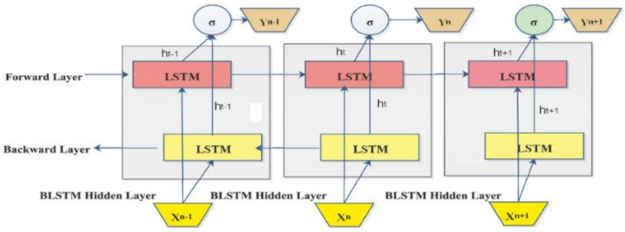


Fig. 3. Bidirectional LSTM Worflow Architecture

The BiLSTM architecture is designed to capture long-range sequential dependencies within EEG segments by processing the input time-series in both the forward and backward temporal directions simultaneously. Standard LSTM units address the vanishing gradient problem of vanilla RNNs through a gating mechanism comprising an input gate i_t, a forget gate f_t, an output gate o_t, and a cell state C_t, governed by the following recurrence relations:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t &= \tanh\,(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t \odot \tanh\,(C_t) \end{aligned}$$

The bidirectional extension augments each LSTM layer with a reverse-direction LSTM that processes the input sequence from the final timestep to the first, producing a backward hidden state $\overleftarrow{h}_t$. The forward and backward hidden states are concatenated at each timestep, yielding a combined representation $[\overrightarrow{h}_t; \overleftarrow{h}_t]$that encodes both past and future temporal context, a property of particular value in EEG analysis, where the onset and cessation of ictal activity are defined by bidirectional contextual patterns.

The BiLSTM network comprises two stacked bidirectional LSTM layers with 128 units each (yielding 256-dimensional concatenated outputs), followed by Dropout regularization with a rate of 0.5 between layers. The recurrent output is passed through two fully connected layers of 128 and 64 units with ReLU activation before the final softmax classification layer. The model is trained using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss.

3. CNN-LSTM Hybrid Architecture

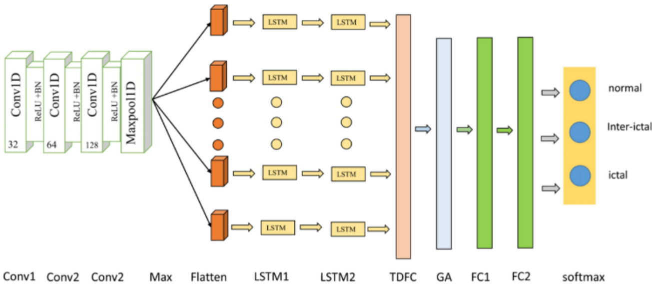


Fig. 4. CNN-BiLSTM Workflow Architecture

The CNN-LSTM Hybrid model is designed to synergistically integrate the spatial feature extraction capabilities of convolutional layers with the sequential temporal modeling capabilities of LSTM units, addressing the complementary limitations of each architecture in isolation. The motivation for this fusion is grounded in the observation that raw EEG segments contain both local morphological features which are transient spike discharges, sharp waves, best captured by convolutional operations, and long-range temporal dynamics, rhythmic oscillations, seizure onset and propagation patterns are best modeled by recurrent units.

The hybrid architecture proceeds in two sequential stages. In the first stage, two one-dimensional convolutional blocks, each comprising a Conv1D layer (64 and 128 filters respectively, kernel size 3), Batch Normalization, ReLU activation, and MaxPooling, they extract a compact sequence of high-level spatial feature maps from the raw input segment. These feature maps, of reduced temporal dimensionality, serve as the sequential input to the second stage. In the second stage, a Bidirectional LSTM layer with 128 units receives the convolutional feature sequence, modeling the temporal evolution and interdependencies of the extracted spatial features across the segment. This two stage design enables the model to learn what features are present at each local temporal position via convolution, and how those features relate and evolve over time via recurrent processing.

The LSTM output is followed by a fully connected layer of 128 units with ReLU activation and Dropout (rate 0.5), before the final softmax output layer. The hybrid model is trained with the Adam optimizer at a learning rate of 0.001. As reported in the experimental results, the CNN-LSTM Hybrid achieves a classification accuracy of 93.1% and an F1 score of 93.1%, confirming that the integration of spatial and temporal learning yields competitive performance, though the pure 1D-CNN marginally outperforms the hybrid on this particular dataset configuration, potentially attributable to the relatively short 512-sample segment length which may limit the sequential context available to the LSTM component.

WORKFLOW

A. Pipeline

The system workflow consists of sequential stages beginning with raw EEG input and ending with classification output:

Raw EEG Signal → Preprocessing → Feature Extraction / Automated Learning → Model Training → Evaluation → Deployment

Raw EEG recordings from the Bonn University dataset (Sets A–E), each containing 4096 samples, are ingested from plain text files and labeled according to their respective classes. Each recording is segmented into 512 sample windows and normalized using Z score normalization, improving data consistency and model convergence.

Two parallel processing paths are then employed. For classical machine learning models, handcrafted features are extracted using Power Spectral Density (PSD) via Welch's method and Discrete Wavelet Transform (DWT) with the db4 wavelet. For deep learning models, the normalized EEG segments are provided directly as input, enabling automatic feature extraction.

The system trains six classifiers within a unified framework: Random Forest, SVM, Decision Tree, 1D-

CNN, BiLSTM, and a CNN–LSTM hybrid model. All models are evaluated using accuracy, precision, recall, F1 score, and confusion matrices on a held out test set.

B. Training Configuration

All deep learning models are trained using a consistent configuration to ensure fair comparison and stable convergence. The Adam optimizer with a learning rate of 0.001 and the categorical cross-entropy loss is used. A batch size of 32 and up to 100 epochs are employed, 20% of the training data being used for validation.

To prevent overfitting, two training callbacks are applied:

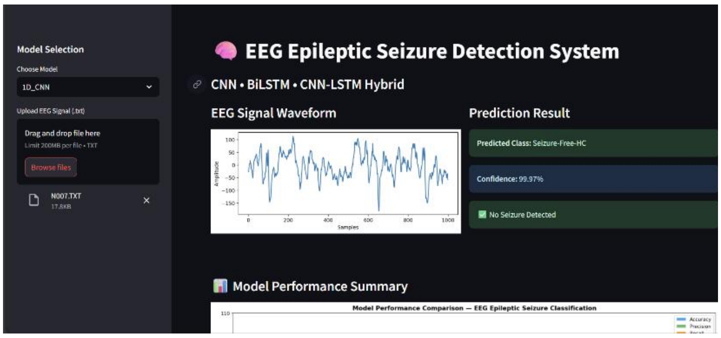
- ReduceLROnPlateau, which reduces the learning rate by a factor of 0.5 when validation loss stagnates for 10 epochs.
- EarlyStopping, which halts training if no validation improvement is observed for 20 epochs and restores the best-performing model weights.

Also, Dropout regularization (rate = 0.5) is added on the dense layers for better generalization.

C. Deployement Architecture

The trained deep learning models are serialized and integrated into a Web application for real time inference. The application accepts raw EEG files, performs preprocessing internally, and outputs the predicted EEG class along with a confidence score and seizure indication. Segment level predictions are aggregated using majority voting to improve robustness against noise and artifacts.

IV. RESULTS AND DISCUSSION



Model Selection

Choose Model: 1D_CNN

Upload EEG Signal (.txt)

Drag and drop file here
Limit 200MB per file • TXT

N007.TXT
17.8KB

EEG Epileptic Seizure Detection System

CNN • BiLSTM • CNN-LSTM Hybrid

EEG Signal Waveform

Prediction Result

Predicted Class: Seizure-Free-HC

Confidence: 99.97%

✓ No Seizure Detected

Model Performance Summary

Model Performance Comparison — EEG Epileptic Seizure Classification

Fig. 5. Output on website classifying into current class

A. Performance Evaluation

Traditional machine learning models show reasonable performance when trained on handcrafted PSD and DWT features; however, they are consistently outperformed by deep learning approaches. This demonstrates the advantage of automated feature learning for complex, non stationary EEG signals.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	84.6	84.9	84.6	84.5
SVM	86.2	86.5	86.2	86.1
1D-CNN	90.8	91.0	90.8	90.8
BiLSTM	91.9	92.1	91.9	91.9
CNN–LSTM Hybrid	93.1	93.3	93.1	93.1

Table.2. Performance Evaluation Table of 5 trained models

B. Confusion Matrix and Training Analysis

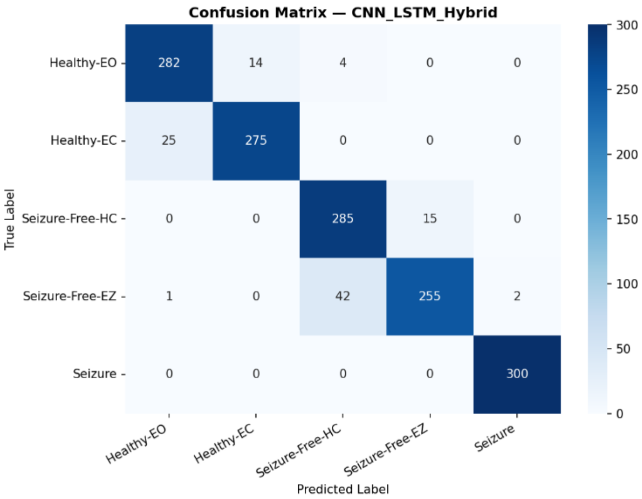


Fig. 6. Confusion Matrix CNN-LSTM-Hybrid

The confusion matrices illustrate that classical models exhibit higher misclassification rates between seizure free and healthy EEG classes. In contrast, deep learning models show improved class separation, with the CNN–LSTM hybrid model achieving the most balanced classification across all five EEG classes.

Training history plots further confirm these findings. The 1D-CNN converges rapidly but exhibits minor fluctuations in validation loss, indicating limited temporal modeling capability. The BiLSTM shows smoother convergence due to its ability to model long term dependencies. The CNN–LSTM hybrid demonstrates both rapid convergence and stable validation performance, indicating superior generalization and reduced overfitting.

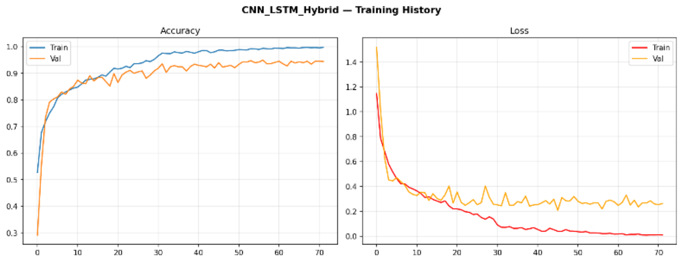


Fig. 7. CNN-LSTM-Hybrid Training Accuracy and Loss

C. Success Analysis of the Hybrid Model

The better accuracy of CNN–LSTM is that it combines the correlation of spatial and temporal features of EEG signals. The CNN extracts localized discriminative features from raw EEG segments, thus the feature extraction process denoises and reduces the dimensionality. These compact feature representations are then passed to the LSTM layers, which capture long term temporal dependencies and sequential patterns inherent in epileptic activity.

Such a trade off in duties can cope with the drawbacks of CNN and LSTM models when used individually. As a result, the CNN–LSTM hybrid attains the best accuracy and F 1 score, and is thus the best model for multi class epileptic seizure detection under the proposed framework.

A. Success Analysis

The results of the experiments strongly indicate that deep learning methods have superior performance for multi class EEG classification. Among the tested models, the CNN–LSTM hybrid model yields the best performance in terms of accuracy and f1-score, which implies that it has the best generalization on the five EEG classes. This progress is mainly attributed to the complementary ability of the proposed hybrid framework along which the well-known CNN layers play a leading role in noise elimination and signal dimensionality reduction by developing local discriminatory features, whereas the LSTM modules represent long term temporal correlations among EEG signals. This spatial temporal fusion leads to more robust discrimination among Seizure, Seizure free and Normal EEG signals than single CNN or BiLSTM models.

B. Failure Analysis

There remains some errors, however, with the most confusion between Healthy-Eyes Open (Set A) and Healthy-Eyes Closed (Set B) signals. These mistakes are due to the very similar neurophysiological patterns of those two classes, which are separated only by fine variations in alpha band activity. A little confusion occurs between seizure free classes (Sets C and D] which is not surprising given the variability and the low amplitude of the interictal EEG 2. Also while conventional classifiers provide faster prediction they are less accurate, since they make prediction based on handcrafted features, thus they can be seen as computational efficient methods with lower classification performance.

V. CONCLUSION

This study presented a comprehensive EEG based epileptic seizure classification framework combining traditional machine learning techniques with advanced deep learning models. Using the Bonn University EEG dataset, a five class classification problem was solved, which distinguishes normal, seizure free and seizure states. Experimental results for deep learning based methods have been showing better performance than traditional classifiers, and CNN–LSTM based hybrid model attains the best performance on average among all evaluation measurement including accuracy, precision, recall and F1-score. The better results on the hybrid model proved that combining convolutional feature extraction with recurrent temporal motion modeling was effective for analyzing complicated EEG signals.

The proposed framework shows promising performance and it can be further enhanced in the future. Testing on larger, more diverse EEG data sets, for example the CHB-MIT scalpeeg database would improve generalization and clinical applicability. Furthermore, real-time hardware implementation of the trained models on hardware platforms such as FPGA-based accelerators or Edge AI devices would enable low-latency seizures detection for continuous patient monitoring. Patient specific adaptation and lightweight model compression methods could be subject of further studies to even further improve performance for clinical use-in real world.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Vishwakarma Institute of Technology, Pune, for providing the resources and environment necessary for this research. The authors also sincerely thank Prof. Amruta Mankawade for her guidance and support, as well as Prof. Shital Dongre, Head of the Department of Artificial Intelligence and Data Science.

REFERENCES

[1] World Health Organization, "Epilepsy: Key facts," WHO, Geneva, Switzerland, Feb. 2023.

[2] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," Physical Review E, vol. 64, no. 6, p. 061907, Nov. 2001.

[3] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," Expert Systems with Applications, vol. 32, no. 4, pp. 1084–1093, May 2007.

[4] E. D. Übeyli, "Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines," Computers in Biology and Medicine, vol. 38, no. 1, pp. 14–22, Jan. 2008.

[5] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in Proc. 27th Int. Conf. Machine Learning (ICML), Haifa, Israel, Jun. 2010, pp. 975–982.

[6] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG analysis of epilepsy: A review," Knowledge-Based Systems, vol. 45, pp. 147–165, Jun. 2013.

[7] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," Computers in Biology and Medicine, vol. 100, pp. 270–278, Sep. 2018.

[8] Q. Yuan, W. Zhou, S. Li, and D. Cai, "Epileptic EEG classification based on extreme learning machine and nonlinear features," Epilepsy Research, vol. 96, no. 1–2, pp. 29–38, Sep. 2011.

[9] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals," Computers in Biology and Medicine, vol. 99, pp. 24–37, Aug. 2018.

[10] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, long short-term memory, fully connected deep neural networks," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, QLD, Australia, Apr. 2015, pp. 4580–4584.

[11] D. Zhang, L. Yang, X. Yuan, and J. Huang, "Epileptic seizure detection using convolutional neural network with transfer learning," in Proc. IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM), Madrid, Spain, Dec. 2018, pp. 1040–1044.

.