International Journal of Emerging Science and Engineering (IJESE)
ISSN: 2319–6378 (Online), Volume-14 Issue-6, May 2026

Hybrid Supervised-Unsupervised Learning Pipeline for EEG Anomaly Detection Using Autoencoders and 1D CNN Models

Onkar Belure, Ayush Awasthi, Ankur Bhutare

Abstract: Seizure detection from electroencephalogram (EEG) signals remains a difficult problem because of the wide variation in patient patterns and the limited amount of labelled data. In this work, we developed a hybrid learning setup that blends a supervised 1D Convolutional Neural Network (CNN) with an unsupervised Autoencoder (AE). The CNN learns to recognise seizure-related patterns from labelled EEG segments, while the AE models typical EEG activity and signals abnormal deviations. We combined their predictions using three ensemble techniques-Soft Weighted Average, Soft Average, and Majority Voting-to stabilize performance and reduce false alarms. Tests on the Turkish Epilepsy EEG Dataset showed that this hybrid approach performed more reliably across patients than either model alone.

Index Terms: EEG, Seizure Detection, Autoencoder, 1D CNN, Hybrid Learning, Ensemble Learning, Anomaly Detection, Biomedical Signal Processing

Nomenclature:
CNN: Convolutional Neural Network
AE: Autoencoder
EEG: Electroencephalogram

I. INTRODUCTION

Epilepsy causes recurrent seizures that disrupt brain function and can severely affect daily life. In hospitals, specialists typically analyse long EEG recordings by hand, a process that can take hours and still relies heavily on personal judgment [1]. This makes automation important. Supervised models, such as CNNs, have achieved encouraging accuracy in seizure detection [4] [7], but they rely on large labelled datasets that are rarely available. In contrast, Autoencoders (AEs) can learn the typical behaviour of brain signals from unlabeled EEG data and later detect deviations as anomalies [3]. However, because AEs do not use seizure labels, they may occasionally mistake motion artefacts or sensor noise for seizure activity [8].

Manuscript received on 07 November 2025 | First Revised Manuscript received on 26 April 2026 | Second Revised Manuscript received on 04 May 2026 | Manuscript Accepted on 15 May 2026 | Manuscript published on 30 May 2026.
*Correspondence Author(s)
Onkar Belure*, Department of Computer Science and Engineering, MIT-ADT University, Pune (Maharashtra), India. Email ID: onkarbelure1905@gmail.com, ORCID ID: 0009-0009-7780-4131
Ayush Awasthi, Department of Computer Science and Engineering, MIT-ADT University, Pune (Maharashtra), India. Email ID: ayushawasthi1125@gmail.com
Ankur Bhutare, Department of Computer Science and Engineering, MIT-ADT University, Pune (Maharashtra), India. Email ID: Bhutareankur@gmail.com

We combine the strengths of both strategies by training a CNN and an AE separately and merging their predictions. This hybrid setup improves reliability by balancing the CNN's precision with the AE's sensitivity.

II. RELATED WORK

Early seizure detection systems relied on manually designed features and statistical classifiers such as SVMs or Random Forest

Forests [1]. The recent shift to deep learning introduced CNNs, RNNs, and Transformer-based models that can learn relevant temporal patterns directly from EEG data [2], [5]. Autoencoders have been used to learn non-seizure signal structures

so that high reconstruction error indicates anomalies [6]. Still, when trained only on clean EEG, AEs often mistake benign irregularities for seizures.

Hybrid learning approaches have emerged to close this gap. Mekruksavanich et al. [3] combined CNNs with attentionenhanced autoencoders to improve robustness. Inspired by this, our study adopts a simpler design and focuses on reliable ensemble fusion using a real EEG dataset.

III. METHODOLOGY

A. Overview

Our workflow has three stages: preprocessing and segmentation; training separate CNN and AE models; and combining their predictions via ensemble fusion.

B. Data Preprocessing

We used the Turkish Epilepsy EEG Dataset [9], containing multi-channel EEG recordings at 500 Hz. Signals were filtered to remove noise and baseline drift, normalized, and divided into overlapping 2 -second windows. The short window length helped preserve seizure onset information while keeping the input manageable.

C. 1D CNN Model

The CNN acts as a supervised binary classifier to separate seizure from non-seizure EEG. For input $x \in \mathrm{R}^{T \times C}$, convolution is given by:

$$y_i = f\ {}_{k=1}^{2K} w_k x_{i+k-1} + b! \qquad (1)$$

where $f(\cdot)$ is the ReLU activation, K is the kernel size, and b is the bias. Output probability is computed as:

$$P_{\mathrm{CNN}} = \sigma(z) = \frac{1}{1+e^{-z}} \qquad (2)$$

We trained the CNN using binary cross-entropy loss and the Adam optimiser (learning rate 0.001,30 epochs). The network contained three convolutional layers followed by dropout and a dense classifier.

D. Autoencoder Model

The AE was trained only on non-seizure EEG data to model normal activity. It consists of an encoder $E(\cdot)$ and decoder $D(\cdot)$:

$$\hat{x} = D\big(E(x)\big). \qquad (3)$$

The training objective minimizes mean squared reconstruction loss:

$$\mathrm{L}_{AE} = \frac{1}{N}\,{}_{i=1}^{N}(x_i - \hat{x}_i)^2 \qquad (4)$$

During testing, segments with high reconstruction error (above threshold τ) are labelled as anomalies. We chose τ empirically based on validation results to reduce false positives.

E. Ensemble Fusion

To merge CNN and AE predictions, we tested three methods:

- Soft Weighted Average:

$$P_{\mathrm{final}} = w_1 P_{\mathrm{CNN}} + w_2 P_{\mathrm{AE}}, w_1 + w_2 = 1. \qquad (5)$$

Weights w_1 and w_2 were adjusted based on the validation F1-score.

- Soft Average:

$$P_{\mathrm{final}} = \frac{P_{\mathrm{CNN}} + P_{\mathrm{AE}}}{2} \qquad (6)$$

- Majority Voting:

$$\text{Label}_{\ \mathrm{final}} = \text{mode}\,\{L_{\mathrm{CNN}}, L_{\mathrm{AE}}\}. \qquad (7)$$

Among these, Majority Voting gave the most stable results.

IV. EXPERIMENTAL SETUP AND RESULTS

A. Hardware and Training Configuration

Experiments were performed on a laptop with an Intel i5 10th Gen CPU and an NVIDIA MX110 GPU. The software environment was TensorFlow 2.x with CUDA acceleration enabled.

B. Training Parameters

i. *C. 1 First Training Phase:*

- Sampling frequency (FS): 500 Hz
- Window length: 2.0 s (1000 samples)
- Step length: 1.0 s (500 samples overlap)
- Channels: All available EEG channels
- K-Folds: 5 (results averaged across folds)
- Batch size: 32
- Epochs: 50
- Seed: 42
- L2 Regularization: $1e^{-4}$
- Dropout rate: 0.3

ii. *C. 2 Fine-Tuning Phase:*

- Validation split: 0.10 (fraction held out for validation)
- Fine-tuning epochs: 20
- Batch size: 32
- Learning rate: $1e^{-4}$
- Seed: 42

C. Model Performance Comparison

Table I: Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
1D CNN	0.63	0.59	0.34	0.44	0.64
Autoencoder	0.63	0.55	0.54	0.55	0.64
Proposed Hybrid	0.67	0.61	0.48	0.53	0.69

D. Comparison with Recent Works

Table II: Comparison with Recent Studies

Study	Year	Method	Accuracy
Mekruksavanich et al. [3]	2025	CNN + Attention AE	0.7
Wang et al. [4]	2023	CNN-LSTM Hybrid	0.66
This Work	2025	1D CNN + AE Ensemble	0.67

E. Analysis

The combination of CNN and AE worked well because each compensated for the other's weaknesses. The CNN learned discriminative seizure patterns, and the AE modelled the overall EEG structure. Together they reduced missed detections with-out, adding many false positives. One practical challenge we faced was tuning AE thresholds: too low caused false positives, while too high missed early seizure onset.

V. CONCLUSION

This study presents a hybrid EEG anomaly detection framework that combines a 1D CNN with an Autoencoder. The hybrid approach provided steadier performance than single models, particularly when labelled data were scarce. In practical terms, this framework could support neurologists by pre-screening long EEG sessions and marking potential seizure segments for review, saving significant time in diagnosis. A limitation observed is the use of fixed ensemble weights; dynamic adjustment based on patient-specific data could further improve accuracy. Future work will extend the system by implementing adaptive weighting, incorporating multimodal physiological data (EEG, ECG, EMG), and deploying it in near-real-time hospital environments. Code and trained models will be made available upon request to support reproducibility.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This

independence ensures that the research is conducted with objectivity and without any external influence.

- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. X. Zhang et al., "A review of epilepsy detection and prediction methods based on EEG signal processing and deep learning," Frontiers in Neuroscience, vol. 18, 2024, DOI: https://doi.org/10.3389/fnins.2024.1468967.
2. Y. Sun et al., "Continuous Seizure Detection Based on Transformer and Long-Term iEEG," IEEE J. Biomed. Health Inform., vol. 26, pp. 54185427, 2022, DOI: http://doi.org/10.1109/JBHI.2022.3199206.
3. S. Mekruksavanich et al., "Epileptic seizure detection in EEG signals via an enhanced hybrid CNN with an integrated attention mechanism," Math. Biosci. Eng., vol. 22, pp. 73-105, 2025, DOI: http://doi.org/10.3934/mbe.%202025004.
4. X. Wang et al., "Automated recognition of epilepsy from EEG signals using a CNN-LSTM hybrid," Scientific Reports, vol. 13, 2023, DOI: http://doi.org/10.1038/s41598-023-41537-z.
5. S. Tan et al., "Automatic detection and prediction of epileptic EEG signals based on nonlinear dynamics and deep learning: a review," Frontiers in Neuroscience, vol. 19, 2025, DOI: http://doi.org/10.3389/fnins.%202025.1630664.
6. J. Xu et al., "EEG-based epileptic seizure detection using deep learning," Neurocomputing, vol. 654, pp. 21-34, 2024, DOI: http://doi.org/10.1016/j.neucom.2024.01.4152.
7. J. Jung et al., "Accuracy of Machine Learning in Detecting Pediatric Epileptic Seizures: A Systematic Review," JMIR Med. Internet Res., vol. 26, p. e55986, 2024, DOI: http://doi.org/10.2196/55986.
8. S. Karthik et al., "Enhanced EEG Signal Processing for Accurate Epileptic Seizure Detection," SN Appl. Sci., vol. 7, 2025, DOI: http://doi.org/10.1007/s42979-025-04148-1.
9. B. Taci, "Turkish Epilepsy EEG Dataset," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/buraktaci/turkish-epilepsy