

# A CNN-BiLSTM Framework for Seizure Detection from Original and Level-Crossing Sampled EEG Signals

Raoua Chouikh

COSIM Laboratory

Sup'Com, University of Carthage  
Tunis, Tunisia

raoua.chouikh@supcom.tn

Asma Maalej

COSIM Laboratory

Sup'Com, University of Carthage  
Tunis, Tunisia

asma.maalej@supcom.tn

Manel Ben Romdhane

COSIM Laboratory

Sup'Com, University of Carthage  
Tunis, Tunisia

manel.benromdhane@supcom.tn

**Abstract**—In hybrid collaborative health networks, transmitting raw electroencephalogram (EEG) signals between nodes creates significant bandwidth and privacy challenges. Signal compression at the acquisition stage can alleviate these issues, but its impact on deep learning-based seizure detection remains poorly understood. This paper presents a comparative study of a convolutional neural network combined with bidirectional long short-term memory (CNN-BiLSTM) architecture with self-attention applied to three signal representations: the original EEG signals of Bonn University, the reconstructed level-crossed signals, and the raw timestamp-free level-crossed samples. A unified pipeline including feature extraction, Bayesian hyperparameter optimization, file-wise 5-fold cross-validation, and file-level majority voting is consistently applied across all signal representations. For the original signals, the model achieves 97.65% accuracy, 97.64% F1-score and 97.40% sensitivity. The timestamp-free and reconstructed signals achieve accuracies of 85.20% and 96.15% at compression ratios of 61.7% and 24.9%, respectively. These results confirm that level-crossing analog-to-digital converter (LC-ADC) compression preserves discriminative information relevant to epileptic EEG classification even after reconstruction, supporting bandwidth-efficient EEG analysis across distributed collaborative networks.

**Keywords**—EEG signal compression, LC-ADC, seizure detection, CNN-BiLSTM, self-attention, Bayesian optimization, collaborative health networks.

## I. INTRODUCTION

Hybrid collaborative health networks enable distributed healthcare delivery across hospitals, clinics, telemedicine platforms, and wearables while respecting local data governance and privacy [1]. However, transmitting electroencephalogram (EEG) signals between nodes remains challenging due to their multidimensional, redundant nature [2]. Compression at the acquisition node offers a solution. Various techniques exist, including lossless methods, wavelet-based coding, and Compressive Sensing [3, 4]. Among these, the level-crossing analog-to-digital converter (LC-ADC) is particularly attractive for biomedical signals as it adapts sampling to signal activity, reducing samples during quiescent periods while preserving rapid variations during seizures [5, 6]. Yet, the impact of

compression-induced distortion on seizure detection performance remains unknown.

While CNN-BiLSTM architectures achieve high performance on uniformly sampled EEG [7, 8], their effectiveness on LC-ADC compressed signals is unexplored. This question is critical for collaborative networks: performance degradation would force a bandwidth-diagnostic trade-off, while preserved diagnostic information would enable bandwidth-efficient collaboration while reducing the volume of raw data transmitted. This paper studies the impact of the LC-ADC compression on CNN-BiLSTM seizure detection compared to uncompressed EEG, framing seizure detection as a binary classification task (epileptic vs. non-epileptic) applied per recording, consistent with prior work on the Bonn dataset. Rather than proposing a new architecture, our contribution is empirical: CNN-BiLSTM, self-attention, and Bayesian optimization are established techniques, but their systematic, controlled combination across three EEG signal representations has not previously been studied. Our contributions are threefold: (1) a unified CNN-BiLSTM pipeline applied identically to all signal types, so performance differences reflect the input representation rather than the model; (2) application to three signal representations (original EEG, reconstructed LC-ADC, raw timestamp-free LC-ADC); (3) a quantitative accuracy-compression trade-off analysis relevant to bandwidth-constrained collaborative networks. Section 2 reviews related work; Section 3 describes the proposed CNN-BiLSTM with LC-ADC pipeline for collaborative seizure detection. Section 4 reports experimental results; Section 5 concludes, discusses limitations, and outlines future work.

## II. RELATED WORK ON EEG COMPRESSION AND CNN-BiLSTM SEIZURE DETECTION

This section reviews the literature relevant to the two core components of the proposed work: EEG signal compression and deep learning-based seizure detection. We first examine deep learning architectures applied to automated seizure detection from EEG signals, with a focus on CNN-BiLSTM models. We then discuss existing EEG compression techniques, with

particular attention to LC-ADC-based approaches and their impact on seizure detection performance.

#### A. Deep Learning for Seizure Detection from Uniformly-Sampled EEG Signals

Deep learning methods, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are widely applied for epileptic seizure detection from uniformly sampled EEG signals [9–12], with CNN-based methods reaching up to 99.2% accuracy [9, 10]. Since CNNs alone cannot capture long-range temporal dependencies, hybrid CNN-LSTM/BiLSTM architectures [7, 8, 11] reach up to 99.39% accuracy [11], and recent work adds self- or multi-head attention [8, 13]. Britto et al. [7] proposed a CNN-BiLSTM framework specifically for seizure detection. All these works and others are based on uniformly sampled EEG signals without compression.

#### B. EEG Signal Compression

Efficient EEG transmission is critical in distributed healthcare. Techniques include wavelet-based methods [14], compressive sensing [3], and LC-ADC, which samples only at amplitude-threshold crossings [5, 6], achieving high compression with minimal distortion. Beyond reconstruction, LC-ADC has been explored for classification via traditional machine learning [15] and spiking neural networks tuned with information-theoretic criteria [16], though both focus on preserved information rather than seizure detection accuracy. Existing studies primarily evaluate reconstruction quality using Mean Squared Error (MSE) or Percent Root-mean-square Difference (PRD), leaving the impact of compression on deep learning models unaddressed — a critical limitation for collaborative health networks.

#### C. Hyperparameter Optimization for CNN-BiLSTM

Hyperparameter choice directly impacts the model's convergence and generalization. Yet manual tuning [7] and grid search [17, 18] quickly become impractical as the number of hyperparameters grows. Bayesian optimization [19] offers an alternative: it builds a surrogate model that iteratively selects the most promising configurations. This yields order-of-magnitude efficiency gains over grid and random search [19], particularly valuable given CNN-BiLSTM's complex hyperparameter interdependencies. Yet it remains underexplored in seizure detection, where most studies still rely on fixed or manually tuned hyperparameters [7, 8]. This work addresses this gap by adopting a unified Bayesian optimization strategy to tune the CNN-BiLSTM hyperparameters consistently across all three signal representations, original EEG signals, LC-ADC reconstructed EEG signals and timestamp-free LC-ADC compressed EEG signals.

Table I summarizes and positions our work relative to the most relevant prior studies.

Table I. Qualitative positioning of the proposed work relative to representative prior studies.

| Reference   | Architecture | Signal | Tuning | Validation   | Compression |
|-------------|--------------|--------|--------|--------------|-------------|
| Acharya [9] | CNN          | EEG    | Fixed  | Random Split | No          |

| Reference             | Architecture         | Signal           | Tuning        | Validation          | Compression |
|-----------------------|----------------------|------------------|---------------|---------------------|-------------|
| Britto et al. [7]     | CNN-BiLSTM           | EEG              | Manual        | 90/10 Split         | No          |
| Liu et al. [8]        | Bilinear CNN-LSTM    | EEG              | N/A           | N/A                 | No          |
| Qaisar & Hussain [15] | ML Ensemble          | LC-ADC based EEG | Grid Search   | 10-Fold CV          | Yes         |
| Safa et al. [16]      | SNN                  | LC-ADC based EEG | Info Criteria | N/A                 | Yes         |
| Proposed work         | CNN-BiLSTM-Attention | LC-ADC based EEG | Bayesian Opt. | File-wise 5-Fold CV | Yes         |

The review reveals three critical gaps: the lack of a controlled comparison across original, reconstructed LC-ADC, and timestamp-free signals; the lack of a unified CNN-BiLSTM pipeline across these representations; and the limited evaluation of compression in terms of downstream classification performance. Thus, in the next subsection, CNN-BiLSTM with LC-ADC pipeline for collaborative seizure detection is proposed.

### III. PROPOSED CNN-BiLSTM WITH LC-ADC PIPELINE FOR COLLABORATIVE SEIZURE DETECTION

This section describes the proposed CNN-BiLSTM with LC-ADC pipeline for collaborative seizure detection. We first present the dataset used for evaluation, then detail the LC-ADC compression scheme, followed by the CNN-BiLSTM architecture with attention mechanism, and finally the Bayesian optimization strategy adopted to tune the acquisition and model hyperparameters across the three signal types considered.

#### A. Bonn University Dataset

All signals come from the Bonn University EEG dataset [20], widely used in epilepsy detection studies. It comprises five subsets (A–E) of 100 single-channel segments each, 23.6 s long (4096 samples, 173.61 Hz), band-pass filtered (0.53–40 Hz) and digitized with a 12-bit ADC: A (normal, eyes open), B (normal, eyes closed), C (interictal, non-epileptic zone), D (interictal, epileptic zone), E (ictal). We use 400 segments for a balanced binary task, A+B (non-epileptic) vs. D+E (epileptic); C is excluded following Andrzejak et al.'s protocol [20], as it does not cleanly fit either class. The epileptic class combines interictal (D) and ictal (E) segments; the task therefore classifies epileptic origin rather than detecting the onset of a seizure within a continuous recording. Figures 1–2 show representative normal (set A) and interictal epileptic (set D) segments. We adopt this dataset as a controlled benchmark for isolating the effect of LC-ADC compression rather than as a state-of-the-art demonstration: its single-channel, single-subject-per-file structure lets the same segments be re-sampled by LC-ADC and compared under identical conditions, harder to control on larger multi-channel datasets. Prior work on this dataset already reports accuracies above 99% on uncompressed signals [9, 11], but no study has evaluated LC-ADC's effect on the accuracy of epilepsy detection.

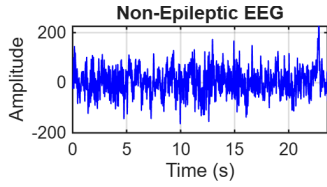


Fig. 1. Representative normal EEG segment (set A - healthy subject, eyes open).

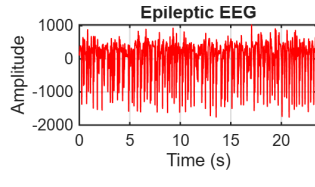


Fig. 2. Representative pathological EEG segment (set D - interictal activity from epileptic zone).

### B. LC-ADC Compression and Reconstruction

Unlike Nyquist-rate ADCs, LC-ADC uses event-driven sampling, generating samples only when the signal crosses predefined amplitude thresholds. Its architecture comprises a comparator, an M-bit digital-to-analog converter (DAC), and a time-tracking module [5, 6]. LC-ADC yields two representations for classification:

1. Reconstructed: timestamps enable linear interpolation reconstruction, preserving both amplitude and temporal information.
2. Timestamp-free: only the sequence of upward/downward crossings is kept; temporal intervals are discarded, and the event stream is converted into fixed-length sequences before feature extraction. For this representation, features are computed from the ordered event sequence rather than physical time, so their index reflects event order rather than elapsed duration.

LC-ADC captures signal transitions only at threshold crossings, generating no samples during quiescent periods — significantly reducing data volume and transmission energy. The two key parameters are quantization levels  $M$  (6–10 bits) and clock frequency  $F_c$  defined via oversampling factor  $OSR_c$  (10–100× relative to 173.61 Hz). Two compression ratio calculations are considered. The first encodes each event using  $M$  bits for amplitude and  $N$  equal to 12 bits for timestamp (matching Bonn dataset's 4096 samples/segment, ensuring lossless encoding). The compression ratio with timestamps (wTS) is then calculated according to equation (1):

$$CR_{wTS} = \left(1 - \frac{(N + M) \times N_{SLC-ADC}}{N_{ADC} \times N_{SADC}}\right) \times 100 \quad (1)$$

where  $N_{SLC-ADC}$  is the number of LC-ADC samples,  $N_{ADC}$  is the original ADC resolution and  $N_{SADC}$  is the number of the original ADC samples.

The second approach ignores the timestamps and encodes only the direction of signal variation using 2 bits: 00 for no variation, 01 for an increase, and 10 for a decrease. Code 11 is reserved and unused, serving as a natural mechanism for detecting transmission errors. In this representation, temporal information is totally discarded. The corresponding compression ratio without timestamps (woTS) is calculated according to equation (2):

$$CR_{woTS} = \left(1 - \frac{2 \times N_{SLC-ADC}}{N_{ADC} \times N_{SADC}}\right) \times 100 \quad (2)$$

Table II presents the compression ratios as a function of  $M$  (6–10) for reconstructed ( $CR_{wTS}$ ) and timestamp-free ( $CR_{woTS}$ ) LC-ADC signals using 2-bit event encoding, according to (1) and (2). These values are the same for all values of  $OSR_c$ .

Table II. Compression Ratio Analysis of (a) LC-ADC Reconstructed Signals ( $CR_{wTS}$ ) and (b) Timestamp-Free LC-ADC Signals ( $CR_{woTS}$ ) using 2-Bit Event Encoding Across Quantization Levels.

| M   | 6     | 7     | 8        | 9         | 10        |
|-----|-------|-------|----------|-----------|-----------|
| (a) | 42.5% | 24.9% | < -94.3% | < -226.9% | < -432.7% |
| (b) | 80.8% | 61.7% | < 26.4%  | < -17.0%  | < -46.5%  |

A positive compression ratio (CR) indicates data reduction; a negative CR indicates data expansion (more events than uniform sampling). Based on these results, only configurations with positive CR ( $M \leq 7$ ) are retained for classification experiments.

### C. Feature Extraction

Feature extraction transforms raw signals into compact, informative representations that improve discriminative power while reducing computational complexity [21]. Statistical, spectral, time-frequency, and dynamical features are extracted uniformly across all three signal types: each 128-sample window yields an 18-dimensional feature vector, and 10 consecutive vectors form an  $18 \times 10$  input sequence — fed to the CNN-BiLSTM in place of the raw waveform, with the CNN operating on this feature sequence and the BiLSTM modeling dependencies across it. Time-domain features include mean, variance, skewness, kurtosis, RMS, energy, zero-crossing rate, and Hjorth parameters (activity, mobility, complexity) [23]. Frequency-domain features estimate band power in Delta, Theta, Alpha, Beta, and Gamma bands plus spectral entropy — established seizure biomarkers [21]. Wavelet-based features (CWT) capture localized time-frequency structure via global energy and its low-, mid-, and high-frequency components, reflecting transient patterns linked to seizure activity. Figures 3 and 4 illustrate representative time-frequency and spectral characteristics of normal and epileptic EEG signals. Compared with normal EEG, epileptic activity exhibits broader spectral variations and increased irregularity across the analyzed frequency range.

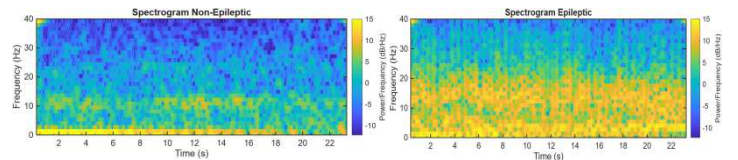


Fig. 3. Spectrogram for (left) normal EEG signal and (right) epileptic EEG signal.

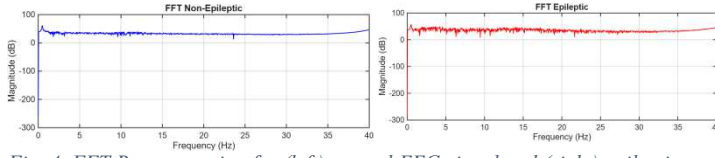


Fig. 4. FFT Representation for (left) normal EEG signal and (right) epileptic EEG signal.

#### D. CNN-BiLSTM Architecture with Self-Attention

A unified deep learning architecture is adopted for all signal types to ensure a fair comparison. The model is composed of three main stages: local feature extraction using convolutional layers, temporal modeling using a bidirectional recurrent network, and attention-based aggregation followed by classification.

##### 1) CNN Configuration

The per-window handcrafted features, summarized in the previous subsection III.C, directly form the model's input sequence: each time step corresponds to one window's 18-dimensional feature vector, and 10 consecutive windows form one input sequence. The CNN module comprises two convolutional layers that learn local relationships among neighboring feature vectors. The first convolutional layer uses 32 filters with a kernel size of 3, followed by batch normalization and ReLU activation. The second convolutional layer uses 64 filters with the same kernel size, followed by batch normalization and ReLU activation. This design progressively transforms the extracted feature representations before temporal modeling by the BiLSTM, increasing the representational capacity of the model while maintaining a compact feature sequence.

##### 2) BiLSTM Configuration

Five key model and training hyperparameters are optimized via Bayesian search. Hidden units are tuned within [64, 128], and the output mode is set to Sequence to support the subsequent self-attention layer. Dropout regularizes the model, while the learning rate (tuned on a logarithmic scale) governs Adam optimizer updates [24], where small deviations from 0.001 meaningfully affect convergence. Mini-batch size balances gradient noise and memory usage. Epoch ranges are adapted to signal type: 10–20 for original signals and 15–25 for compressed signals to ensure convergence. These hyperparameters are jointly optimized in a unified Bayesian strategy, ensuring fair comparison across all signal representations.

##### 3) Self-Attention Mechanism

The BiLSTM outputs are passed through a self-attention layer with a single attention head. This layer assigns adaptive importance weights to each time step in the sequence, allowing the model to focus on the most informative EEG segments [8]. For epileptic EEG classification, this mechanism can emphasize the most informative temporal patterns within the input sequence. The attention output is aggregated via global average pooling into a fixed-length representation for classification. Overall, the CNN layers provide robust local feature extraction, the BiLSTM captures long-range bidirectional dependencies, and self-attention adaptively weights the most informative time steps while keeping the model compact for distributed deployment. The same network architecture and optimization procedure are applied to all signal types (original EEG, LC-

ADC reconstructed EEG, and timestamp-free LC-ADC), while hyperparameters are independently tuned within the predefined search space for each; this ensures that performance differences reflect the input representation rather than the model. All hyperparameters are tuned using a unified Bayesian optimization framework.

#### E. Optimization Strategy

The performance of deep learning models is highly sensitive to hyperparameter selection. Instead of relying on grid search or random search, Bayesian optimization is employed due to its superior efficiency and convergence properties [19]. This approach constructs a probabilistic surrogate model of the objective function and iteratively selects promising hyperparameter configurations. The optimization process balances: exploration: investigating new regions of the search space, and exploitation: refining regions known to yield high performance. The acquisition function is Expected Improvement (EI) [22], prioritizing configurations likely to outperform the current best result. A common Bayesian optimization procedure is used for all signal types (original EEG, reconstructed LC-ADC, timestamp-free LC-ADC), with the epoch range adjusted for compressed signals.

Table III presents the hyperparameter search space for Bayesian optimization, including architectural and training parameters of the CNN-BiLSTM model with their ranges and descriptions.

Table III. Hyperparameter Search Space.

| Parameter       | Range                            | Description             |
|-----------------|----------------------------------|-------------------------|
| Hidden Units    | 64–128                           | BiLSTM capacity         |
| Dropout Rate    | 0.20–0.35                        | Regularization strength |
| Learning Rate   | $10^{-4}$ – $3.5 \times 10^{-3}$ | Step size               |
| Mini-Batch Size | 16–64                            | Batch size              |
| Epochs          | [10–20] / [15–25]                | Training duration       |

For compressed signals, we use a slightly larger epoch range to compensate for compression-induced variability. The optimization objective is maximizing the F1-score rather than raw accuracy. F1-score was chosen because it jointly accounts for precision and recall, which is directly relevant to seizure detection where both missed seizures and false alarms carry a cost. Clinically, this also reflects the dual cost of seizure detection errors: a missed seizure (low recall) delays treatment, while a false alarm (low precision) causes unnecessary clinical follow-up; optimizing F1-score avoids favoring one failure mode over the other by default [8]. To ensure robustness, we evaluate performance using file-wise 5-fold cross-validation (CV), training and testing each configuration across all five folds [18]. The same five folds were used both for Bayesian hyperparameter selection and for the final reported evaluation. Bayesian optimization provides three key advantages: (1) fewer training runs than exhaustive search, (2) equally optimized configurations across all signal types, and (3) automatic

hyperparameter adjustment for different signal representations. The next section presents the experimental results.

#### IV. EXPERIMENTAL RESULTS

The following subsections present results for each signal type and assess the effects of LC-ADC compression and Bayesian optimization on classification performance.

##### A. Hyperparameter Optimization Results

Bayesian optimization is employed to tune the hyperparameters of the CNN-BiLSTM model.

###### 1) Original Signals and LC-ADC Compressed Signals

###### a) Bonn University EEG signals

Bayesian optimization selects 97 hidden units, a dropout rate of 0.33, a learning rate of 0.0009, and a mini-batch size of 42. Averaged over 5 training seeds, the model reaches 97.65% accuracy, 97.64% F1-score, 97.40% sensitivity, and 97.90% specificity. The high sensitivity indicates strong detection of the epileptic class. Figure 5 shows the Receiver Operating Characteristic (ROC) curve with an Area Under the Curve (AUC) of 0.9968, confirming excellent separation between normal and epileptic signals.

###### b) Reconstructed LC-ADC signals ( $M = 7$ , $OSR_c = 20$ , $CR = 24.9\%$ )

The model maintains strong performance, achieving 96.15% accuracy, 96.12% F1-score, 95.60% sensitivity, and 96.70% specificity. The optimized configuration consists of 87 hidden units, a dropout rate of 0.2051, a learning rate of 0.0035, and a mini-batch size of 57. The optimized configuration uses a higher learning rate and a larger mini-batch size than the original-signal configuration. As shown in Figure 5, the AUC of 0.9946 confirms that high discriminative ability is largely preserved despite the 24.9% compression ratio.

###### c) Timestamp-free LC-ADC signals ( $M = 7$ , $OSR_c = 10$ , $CR = 61.7\%$ )

In the absence of timing information, the model achieves 85.20% accuracy, 84.10% F1-score, 78.30% sensitivity, and 92.10% specificity, with 92 hidden units, a dropout rate of 0.2022, a mini-batch size of 32, and 18 epochs. These results indicate that the signal retains discriminative information even without timestamps. As shown in Figure 5, the AUC of 0.9330 indicates solid overall discriminability despite the complete absence of temporal information.

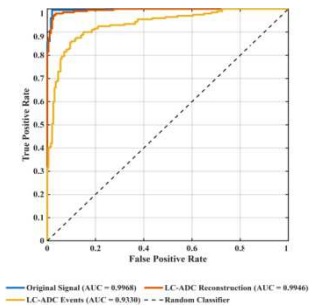


Fig. 5. ROC curves for original EEG, reconstructed LC-ADC, and timestamp-free LC-ADC representations.

##### B. Classification Performance Comparison

Performance is evaluated at the file level using majority voting. Each file is transformed into fixed-length feature sequences using the windowing and feature-extraction procedure, and each sequence is classified independently. The file's final label is the majority vote across all its sequence-level predictions. This evaluation level is more challenging than segment-level assessment since the model must maintain consistent predictions across all segments of a recording, and it is more clinically relevant as in practice a diagnosis is established from a complete EEG recording rather than from isolated signal windows. Table IV summarizes the performance of all signal representations.

Table IV. Performance across signal types

| Metric            | Original   | Reconstructed | Timestamp-Free |
|-------------------|------------|---------------|----------------|
| Accuracy          | 97.65±0.38 | 96.15±0.68    | 85.20±1.37     |
| F1-Score          | 97.64±0.38 | 96.12±0.73    | 84.10±1.58     |
| Sensitivity       | 97.40±0.82 | 95.60±1.88    | 78.30±2.08     |
| Specificity       | 97.90±0.65 | 96.70±0.57    | 92.10±1.29     |
| Compression Ratio | NA         | 24.9          | 61.7           |

File-level evaluation is more clinically realistic than segment-level assessment. Averaged over 5 seeds, our CNN-BiLSTM with self-attention achieves 97.65% accuracy on original EEG, dropping modestly to 96.15% with LC-ADC reconstruction ( $CR=24.9\%$ ) — confirming that discriminative features survive reconstruction — and to 85.20% for timestamp-free signals ( $CR=61.7\%$ ), a 10.95-point gap reflecting the cost of discarding timestamps. The timestamp-free representation provides the highest compression ratio but incurs a substantial sensitivity loss (97.40% vs. 78.30%). The optimized configurations also differed across representations: the original signals used a higher dropout rate (dropout 0.33 vs. 0.205/0.202), whereas the reconstructed signals used a larger mini-batch size (57 vs. 42/32). The timestamp-free representation remains the most bandwidth-efficient option. To assess the contribution of individual architectural components, Table V presents an ablation study across the three signal representations.

Table V. Ablation study (single seed, fixed fold partition) for original, reconstructed, and timestamp-free signals.

| Variant              | Accuracy |        |         | Sensitivity |        |         |
|----------------------|----------|--------|---------|-------------|--------|---------|
|                      | Orig.    | Recon. | TS-Free | Orig.       | Recon. | TS-Free |
| Full (proposed)      | 97.75    | 96.50  | 85.00   | 98.00       | 95.50  | 78.50   |
| No self-attention    | 98.00    | 95.75  | 84.00   | 98.50       | 94.50  | 75.50   |
| Uni-directional LSTM | 96.75    | 96.75  | 82.50   | 96.50       | 97.00  | 76.00   |
| Single CNN layer     | 97.75    | 94.75  | 83.25   | 98.00       | 93.50  | 73.50   |

Table V ablates the full model (same hyperparameters, fold partition, seed). Self-attention provides small gains for reconstructed and timestamp-free signals, while slightly reducing accuracy on the original signals, suggesting it is not the primary driver of accuracy in this setting. Bidirectionality helps substantially on original and timestamp-free signals (+1.00 and

+2.50 points respectively) but not on reconstructed signals, while the second CNN layer matters equally for both compressed representations (+1.75 points each if removed) but not for the original signal — indicating that compressed inputs specifically benefit from the additional convolutional capacity.

## V. CONCLUSION

This paper compares epileptic seizure detection across three EEG representations using the same CNN-BiLSTM self-attention architecture: original, reconstructed LC-ADC (CR = 24.9%), and timestamp-free LC-ADC (CR = 61.7%), via a unified pipeline (feature extraction, Bayesian optimization, 5-fold cross-validation, file-level voting). Averaged over 5 seeds, the model achieves 97.65%, 96.15%, and 85.20% accuracy respectively. Even without timestamps, the signal retains solid discriminative power (78.30% sensitivity, 92.10% specificity). Although sensitivity remains substantially lower than for the original signals, this trade-off may be relevant in bandwidth-constrained applications. These results should be interpreted in light of several limitations. The ablation study shows self-attention and CNN depth contribute unevenly across representations rather than universally, comparisons to prior work (Table I) remain qualitative rather than head-to-head on a shared split, and the Bonn dataset's small, single-channel format limits generalization to larger, multi-channel recordings. Nonetheless, this CNN-BiLSTM framework provides a basis for bandwidth-efficient EEG analysis in collaborative health networks. Future work includes head-to-head comparison with prior methods, validation on multi-channel datasets, federated learning, and pre-ictal prediction.

## ACKNOWLEDGMENT

The authors thank the COSIM Laboratory, Sup'Com, University of Carthage, for supporting this research.

## REFERENCES

- [1] G. R. Abijitha, S. Jothi, and A. Chandrasekar, "Federated deep learning model for epilepsy seizure detection using electroencephalogram (EEG) signal," *Neurological Research*, early access, Sep. 2025, doi: [10.1080/01616412.2025.2555516](https://doi.org/10.1080/01616412.2025.2555516).
- [2] B. H. Brinkmann, M. R. Bower, K. A. Stengel, and G. A. Worrell, "Large-scale electrophysiology: Acquisition, compression, encryption, and storage of big data," *J. Neurosci. Methods*, vol. 180, no. 1, pp. 185–192, 2009, doi: [10.1016/j.jneumeth.2009.03.022](https://doi.org/10.1016/j.jneumeth.2009.03.022).
- [3] S. Aviyente, "Compressed sensing framework for EEG compression," in *Proc. IEEE/SP 14th Workshop Stat. Signal Process. (SSP)*, Madison, WI, USA, Aug. 2007, pp. 181–184.
- [4] N. Sriraam, "A high-performance lossless compression scheme for EEG signals using wavelet transform and neural network predictors," *Int. J. Telemed. Appl.*, vol. 2012, Art. no. 302581, 2012, doi: [10.1155/2012/302581](https://doi.org/10.1155/2012/302581).
- [5] T. Marisa, T. Niederhauser, A. Haeberlin, R. A. Wildhaber, R. Vogel, J. Goette, *et al.* "Pseudo asynchronous level crossing ADC for ECG signal acquisition," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 2, pp. 267–278, Apr. 2017, doi: [10.1109/TBCAS.2016.2619858](https://doi.org/10.1109/TBCAS.2016.2619858).
- [6] S. M. Qaisar, "Efficient EEG signal compression and reconstruction using level-crossing sampling," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, 2019, pp. 1–5.
- [7] K. R. Aravind Britto, S. Srinivasan, S. K. Mathivanan, M. Venkatesan, M. B. Benjula Anbu Malar, S. Mallik, and H. Qin, "A multi-dimensional hybrid CNN-BiLSTM framework for epileptic seizure detection using electroencephalogram signal scrutiny," *Syst. Soft Comput.*, vol. 5, Art. no. 200062, 2023, doi: [10.1016/j.sasc.2023.200062](https://doi.org/10.1016/j.sasc.2023.200062).
- [8] S. Liu, J. Wang, and Y. Chen, "A multi-dimensional hybrid bilinear CNN-LSTM model for epileptic seizure detection and prediction," *J. Neural Eng.*, vol. 21, Art. no. 066045, 2024, doi: [10.1088/1741-2552/ada0e5](https://doi.org/10.1088/1741-2552/ada0e5).
- [9] U. R. Acharya *et al.*, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, 2018, doi: [10.1016/j.combiomed.2017.09.017](https://doi.org/10.1016/j.combiomed.2017.09.017).
- [10] I. Ullah, M. Hussain, E.-U.-H. Qazi, and H. Aboalsamh, "An automated system for epilepsy detection using EEG brain signals based on deep learning approach," *Expert Syst. Appl.*, vol. 107, pp. 61–71, 2018, doi: [10.1016/j.eswa.2018.04.021](https://doi.org/10.1016/j.eswa.2018.04.021).
- [11] G. Xu *et al.*, "A one-dimensional CNN-LSTM model for epileptic seizure recognition using EEG signal analysis," *Front. Neurosci.*, vol. 14, Art. no. 578126, 2020, doi: [10.3389/fnins.2020.578126](https://doi.org/10.3389/fnins.2020.578126).
- [12] E. Tuncer and E. Doğru Bolat, "Classification of epileptic seizures from electroencephalogram (EEG) data using bidirectional short-term memory (Bi-LSTM) network architecture," *Biomed. Signal Process. Control*, vol. 73, Art. no. 103462, Mar. 2022, doi: [10.1016/j.bspc.2021.103462](https://doi.org/10.1016/j.bspc.2021.103462).
- [13] A. Guhdar, R. J. Mstafa, and A. O. Mohammed, "Hybrid deep learning model for epileptic seizure classification using 1D-CNN with multi-head attention mechanism," *\*Biomed. Signal Process. Control\**, vol. 112, Art. no. 108495, 2026, doi: [10.1016/j.bspc.2025.108495](https://doi.org/10.1016/j.bspc.2025.108495).
- [14] K. Srinivasan, J. Dauwels, and M. R. Reddy, "Multichannel EEG compression: Wavelet-based image and volumetric coding approach," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 1, pp. 113–120, Jan. 2013, doi: [10.1109/TITB.2012.2194298](https://doi.org/10.1109/TITB.2012.2194298).
- [15] S. F. Hussain and S. M. Qaisar, "Epileptic seizure classification using level-crossing EEG sampling and ensemble of sub-problems classifier," *Expert Syst. Appl.*, vol. 191, Art. no. 116356, Apr. 2022, doi: [10.1016/j.eswa.2021.116356](https://doi.org/10.1016/j.eswa.2021.116356).
- [16] A. Safa, J. Van Assche, C. Frenkel, A. Bourdoux, F. Cathoor, and G. Gielen, "Exploring information-theoretic criteria to accelerate the tuning of neuromorphic level-crossing ADCs," in *\*Proc. 2023 Annual Neuro-Inspired Computational Elements Conf. (NICE 2023)\**, San Antonio, TX, USA, Apr. 2023, pp. 63–70, doi: [10.1145/3584954.3584994](https://doi.org/10.1145/3584954.3584994).
- [17] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 24, pp. 2546–2554, 2011.
- [18] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, 2012.
- [19] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 25, pp. 2951–2959, 2012.
- [20] R. G. Andrzejak *et al.*, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity," *Phys. Rev. E*, vol. 64, no. 6, Art. no. 061907, 2001, doi: [10.1103/PhysRevE.64.061907](https://doi.org/10.1103/PhysRevE.64.061907).
- [21] X. Zhang, X. Zhang, Q. Huang, and F. Chen, "A review of epilepsy detection and prediction methods based on EEG signal processing and deep learning," *\*Front. Neurosci.\**, vol. 18, Art. no. 1468967, 2024, doi: [10.3389/fnins.2024.1468967](https://doi.org/10.3389/fnins.2024.1468967).
- [22] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *\*J. Global Optim.\**, vol. 13, no. 4, pp. 455–492, 1998, doi: [10.1023/A:1008306431147](https://doi.org/10.1023/A:1008306431147).
- [23] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.
- [24] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015. [Online]. Available: <https://arxiv.org/abs/1412.6980>