A COMPARATIVE STUDY OF EEG SIGNAL PROCESSING TECHNIQUES FOR SEIZURE DETECTION AND PREDICTION USING DEEP LEARNING

Alexe CIUREA [1], Cristina MANOILĂ [2], Bogdan IONESCU [3]

This paper presents a comparative study of EEG signal processing techniques for seizure detection and prediction using the CHB-MIT dataset. Specifically, it compares envelope-based techniques: Hilbert transform, root mean square (RMS), logarithmic power, temporal derivatives, and bandpass filtering. A lightweight convolutional neural network (CNN) is employed to assess performance across EEG clip lengths of 1 and 5 seconds. Preprocessing with bandpass filtering and temporal derivatives shows more stable validation performance and outperforms raw EEG input. Among all techniques, the combination of bandpass filtering and first-order derivatives delivers the best performance.

Keywords: EEG, Hilbert, RMS, log-power, envelope, derivative, bandpass, seizure detection and prediction, CNN, CHB-MIT, preprocessing

1. Introduction

Electroencephalography (EEG) is a widely used technique for monitoring brain activity, commonly applied in the detection of epileptic seizures by the clinician [1]. With the advancement of machine learning, particularly deep learning models such as Convolutional Neural Networks (CNNs), raw EEG signals have been increasingly utilized for classification tasks. However, raw EEG data is often noisy and varies significantly across patients, presenting challenges for model validation and generalization [2].

To address these challenges, various envelop-based techniques have been proposed. For instance, Hilbert Transform-based envelope [3] is used for its ability to represent amplitude modulations and reduce phase noise in the signals. Ujma et al. [4] demonstrated the utility of the instantaneous amplitude obtained via the Hilbert transform to characterize rhythmic amplitude changes in sleep EEG. Root-mean-square (RMS) energy is applied to smooth power fluctuations over time [5].

[1] Doctoral School of Electronics, Telecommunications & Information Technology, The National University of Science and Technology POLITEHNICA of Bucharest, Romania, e-mail: alexe.ciurea@gmail.com
[2] Independent researcher, e-mail: cristinamanoila28@gmail.com
[3] Doctoral School of Electronics, Telecommunications & Information Technology, The National University of Science and Technology POLITEHNICA of Bucharest, Romania, e-mail: bogdanlapi@gmail.com

Computing raw signal power was used for EEG-related event synchronization and desynchronization quantification, where it highlighted event signals of high amplitude over background EEG [6]. Additionally, the logarithm function has been used to compress the dynamic range [7].

In the past three years, models with minimum preprocessing or feature extraction techniques have been evaluated for detecting and predicting seizures, particularly on the CHB-MIT dataset [8]. For example, Cao et al. [9] developed a feature fusion methodology that applies the discrete wavelet transform (DWT) to decompose the raw EEG signal from which the features are extracted and selected. Each decomposition applies a bandpass filter. After decomposition, the selection of the features is performed using a Support Vector Machine-Recursive Feature Elimination (SVM-RFE) model, and then these are fed into classifiers. Their hybrid CNN and bidirectional-long-short-term-memory (Bi-LSTM) model achieves an accuracy of 98.43% on patient-specific seizure detection. However, this approach mainly relies on wavelet-based preprocessing and complex neural architecture, yet it is unclear on the windowing method and validation techniques.

Xiang et al. [10] introduced a synchronization-based graph spatial-temporal attention network (SGSTAN) using Graph Attention Layers so the spatial features can be captured automatically, followed by Transformer layers for temporal features. The method was evaluated on 19 CHB-MIT patients, using 18 common EEG channels and an overlapping sliding window to extract additional pre-ictal segments. They obtained 94.76% accuracy for a 1-second and 98.20% for a 5-second window for inter-ictal and pre-ictal predictions, after removing the direct component at 0 Hz and frequency components ranging from 57 to 63 Hz. The model comes with a high computational complexity and cost, making the signal processing less clear and lacking generalization due to the overlapping technique applied [11].

Shen et al. [12] proposed an approach using short-time Fourier transform (STFT) and Google-Net CNN for real-time detection of the seizure onset using 16 selected CHB-MIT patients and 6 EEG channels. To preprocess the data, they applied a bandpass filter using a 6^{th}-order Butterworth algorithm, a sliding window with a duration of 1.35 seconds, and a 1-second overlap. The STFT output is used to generate time-frequency representations that are fed into the CNN. The paper reports an average accuracy of 97.74% and a sensitivity of 98.90% across all patients. Although their results look promising, the method comes with the risk of data leakage due to overlapping, and, similar to Xiang et al., compromising generalization.

Deng et al. [13] also used overlapping at different windows, 1-second or 0.5-second, to segment samples of seizure and non-seizure for 23 CHB-MIT patients. They selected 400 seconds of seizure and 3600 seconds of non-seizure samples. Their preprocessing workflow applies a bandpass filter to the EEG signal ranging from 1 to 40 Hz and selects 16 channels, which reconstruct the brain

connectivity graphs. These, along with the raw signals, are fed into a dual-branch Deepwalk-Transformer Spatiotemporal fusion network (Deepwalk-TS) for seizure detection. One of the branches handles reconstruction of the channel graph, and the other performs feature extraction, and a 5-fold cross-validation is employed. The resultant accuracy is 99.54%, but similarly to Shen et al., this result may be biased due to the overlapping windows and incurs high computational costs.

Malakouti et al. [14] presented a lightweight seizure detection model by extracting time frequency features using five-level DWT to decompose the EEG signal and 0.5-45 Hz Butterworth band pass filtering on the Bonn EEG dataset [15]. They used adaptive classifiers, which include Adaptive Resonance Theory (ART) and Learning Vector Quantization (LVQ), and enhanced them through grid search and ensemble methods for ictal and interictal tasks. The best performance obtained is 0.85 AUC using the ART classifier. However, the paper does not specify any window size or overlap, if any.

Karthik et al. [16] also used the Bonn dataset, aimed to improve seizure detection; seizure and non-seizure, by applying an Improved Feature Space method (ICFS) with DWT to extract features, which are fed into various SVMs. They used different sub-bands from 2 to 150 Hz and obtained an average accuracy of 97% during validation, though no details are provided regarding the clip length, overlap, or validation technique, if any.

On the TUSZ dataset [17], Yan et al. [18] demonstrated a dynamic temporal-spatial graph attention network (DTS-GAN) achieving up to 91% accuracy in the classification of seven seizure types. The EEG signals are segmented using a time window overlap ratio of 0.5 and 0.75, applied to different window lengths: 1, 2, 4, 8, and 16 seconds. The first step is to perform a Fast Fourier Transform (FFT) on each segment. Next, this is fed into a Spatial module consisting of connectivity map blocks and attentive graph convolution, and into a Long Short-Term Memory (LSTM) model. The two outputs from the LSTM and spatial module are passed to a fully connected layer to generate the predictions. Their method shows that adapting to spatial patterns boosts performance compared to the baseline graph attention network, and does not seem to show a significant improvement. Additionally, the work does not detail its validation strategy, making it difficult to assess the implementation.

Rukhsar et al. [19] processed EEG signals using bandpass filtering between 0 and 128 Hz and a sampling rate of 256 Hz from the CHB-MIT dataset. They excluded the last patient, chb24. Their proposal is a Lightweight Convolution Transformer (LCT) able to learn spatial and temporal information within the EEG signals to detect ictal and interictal events, using cross-patient data for training and validation. Features are extracted at the output of the CNN, then fed into a multi-head attention block, followed by a Multi-Layer Perceptron (MLP). The model

achieves an accuracy of 96.31% for a 0.5-second clip and 95.97% for a 1-second clip length with a 25% overlap.

Chung et al. [20] employed a compact CNN, single channel on CHB-MIT, supporting minimal processing. They pre-selected 13 out of 24 patients and divided the ictal and interictal periods into 4-second-long clips using a 4-second overlapping window with a 1-second step size. As a preprocessing technique, they applied a bandpass filter between 1 and 30 Hz, covering alpha (1-4 Hz) and beta (13-30 Hz) frequency bands. Features are extracted by concatenating the output of two-dimensional CNNs. The best accuracy is achieved using the 18-channel detector, reaching an average accuracy of 98.47%, evaluated using k-fold cross-validation for training and validation. However, the use of an overlapping window may limit real-world applicability when overlapped data is used for testing.

Prakash et al. [21] combined a CNN and Bi-LSTM mechanism capturing local and long-term temporal dependencies. The output features from the two methods are concatenated and fed into a dense layer, so the number of parameters can be reduced to improve cost efficiency. Bandpass filtering within 0.1 and 100 Hz is used to reduce the noise. An additional algorithm, like Pan-Tompkins, is utilized to identify the locations of the R-peaks in the signals. The best accuracy obtained is 99.21% for MIT-BIH A [22], which is publicly available, and 96.17% for the INCART dataset, using 10-fold cross-validation. However, the paper does not specify the use of an overlapping window.

Esmaeilpour et al. [23] presented a method that can identify the preictal region that occurs post the onset of the seizures. The preprocessing involves the extraction of the signals, which are divided into segments of non-overlapping 30 seconds, and a short-term Fourier transformation is applied so the signals can be fed into the CNN. The features are then extracted from the network and fed to a fully connected layer and two classification models, an SVM and a Random Forest. The output will be based on the maximum voting to classify the preictal from interictal segments. The best sensitivity of 90.76% is obtained in the SVM model using CHB-MIT. This method offers reliable prediction, though its long input EEG segments limit temporal precision, and it is limited to a single metric.

While techniques such as filtering [24] and envelope extraction have been explored, there is limited work that systematically compares them within a unified pipeline, on the same dataset, using the same convolutional neural network (CNN) architecture. Furthermore, previous approaches do not evaluate how much methods affect both seizure detection and prediction, nor do they offer configurable preprocessing for adapting to both patient-specific and global data.

To address these gaps, this paper evaluates five signal processing strategies: bandpass filtering, temporal derivatives, Hilbert envelope, RMS envelope, and log-power envelope. These methods are tested on the CHB-MIT Scalp EEG database

using 1-second and 5-second clips. Performance is assessed using our previously proposed lightweight CNN with an attention mechanism [25].

Our goal is to assess how each method influences the model's stability and classification accuracy across both seizure detection and prediction tasks, using non-overlapping segments to ensure realistic evaluation within shorter EEG clips.

2. Dataset

The CHB-MIT Scalp EEG Database consists of EEG recordings from 24 pediatric subjects. Each subject's data includes multiple sessions containing seizure and non-seizure intervals, recorded using 22 EEG channels at a 256 Hz sampling rate. For this study, we segmented recordings into clips of 1 s and 5 s and used a combination of detection (seizure vs non-seizure) and early prediction tasks. Only segments with clear onset annotations were included.

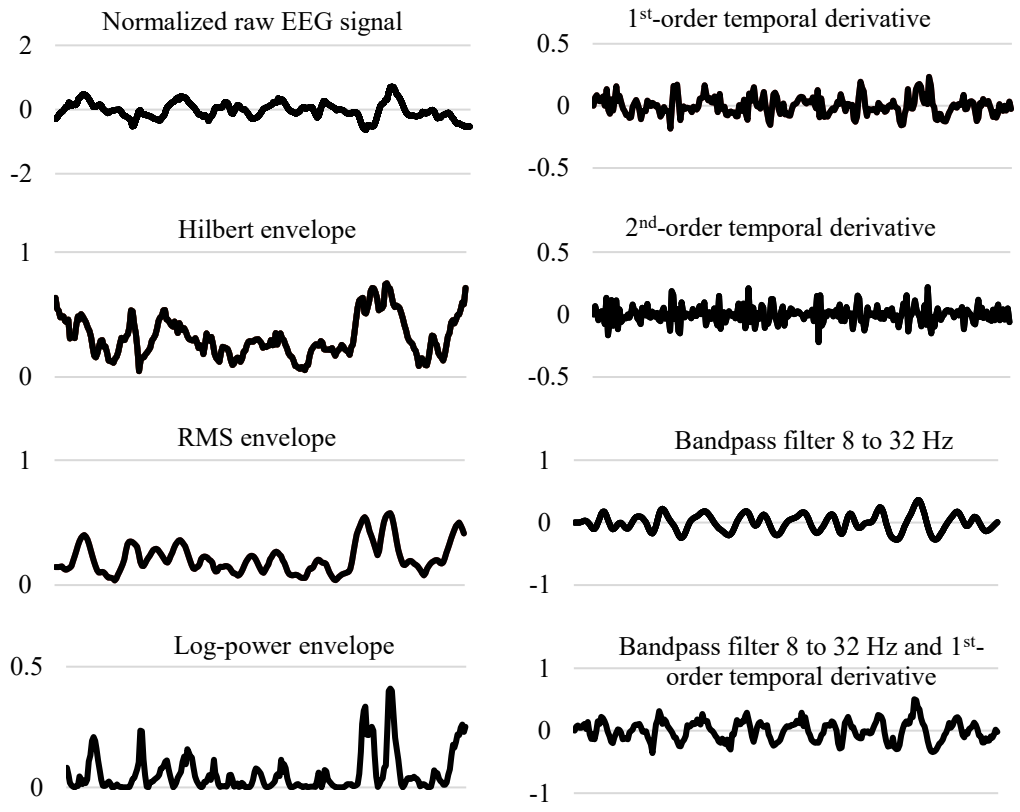

Fig. 1. Comparison of different preprocessing techniques applied to a 1-second-long clip 0, channel 0 of patient chb01, labeled as seizure. The x-axis represents the number of samples (0 to 255, not shown), and the y-axis represents the amplitude.

3. Methods

We represent the input EEG data as a three-dimensional array $X \in \mathbb{R}^{N \times C \times T}$, where N is the number of samples, C is the number of EEG channels, and T is the number of time points. This global representation includes data from all patients. For patient-specific processing, the same data structure is used, but applied individually to each subject.

3.1. Hilbert envelope

The Hilbert transform extracts the analytic signal, which contains both amplitude and phase information of a real-valued EEG signal. The analytic signal $z(t)$ is defined as:

$$z(t) = x(t) + j \times \mathcal{H}\{x(t)\} \tag{1}$$

where $x(t)$ is the real-valued EEG signal, $\mathcal{H}\{x(t)\}$ is the Hilbert transform of $x(t)$ and j is the imaginary unit.
The envelope (or instantaneous amplitude), as seen in Fig. 1, is then computed as the magnitude of this analytic signal:

$$E_{Hilbert}(t) = |z(t)| = \sqrt{x(t)^2 + \mathcal{H}\{x(t)\}^2} \tag{2}$$

3.2. RMS envelope

The Root Mean Square (RMS) envelope estimates the short-term signal using a sliding window average of the squared amplitude. This corresponds to the following formula:

$$E_{RMS}(t) = \sqrt{\frac{1}{N} \times \sum_{i=t-\frac{N}{2}}^{t+\frac{N}{2}} x(i)^2} \tag{3}$$

where $x(i)$ is the EEG signal at time i, N is the window size, and $E_{RMS}(t)$ is the RMS envelope at time t. A window size of 8 samples was chosen empirically as a balance between temporal precision and smoothing.

3.3. Log-power envelope

This method applies a logarithmic transformation to the squared value of the input and corresponds to the following formula:

$$E_{log-power}(t) = log[x(t)^2 + \epsilon] \tag{4}$$

where $x(t)$ is the EEG signal at time t, where ϵ set equal to 1 is a constant to avoid computing large negative numbers, and $E_{log-power}(t)$ is the log-power envelope at time t, with the aim of highlighting seizure features.

3.4. Temporal derivative

The temporal derivative of the raw EEG signal is computed to emphasize dynamic features such as sudden transitions, slope changes, and sharp bursts of activity [26]. First, second, and third order derivatives are calculated using differences along the temporal axis. The first-order derivative captures the signal's rate of change, the second-order highlights acceleration, and the third-order emphasizes curvature. These features are particularly useful for enhancing the visibility of rapid changes often associated with seizure onset.

For example, the first order derivative is defined as:

$$E_{derivative}(t) = f' = \frac{dx(t)}{dt} = x(t) - x(t-1) \tag{5}$$

where $x(t)$ is the signal at time t. This difference estimates the rate of change between consecutive time points in the discretely sampled EEG signal.

3.5. Bandpass filtering

A 4th-order Butterworth bandpass filter is applied to remove both low-frequency and high-frequency noise while preserving the shape of the EEG waveform. The filter is designed using forward and backward filtering to avoid phase distortion. Both the cutoff frequencies and filter order are configurable, allowing for either global settings (applied uniformly across all patients) or patient-specific configurations. This method ensures that only relevant frequency bands are retained for further analysis.

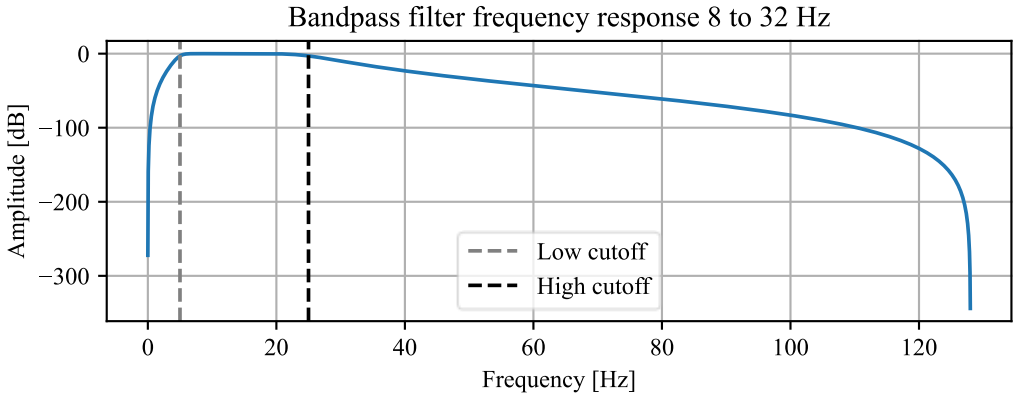

Fig. 2. Frequency response of a 4th-order Butterworth bandpass filter with cutoff frequencies at 8 Hz and 32 Hz (sampling rate = 256 Hz)

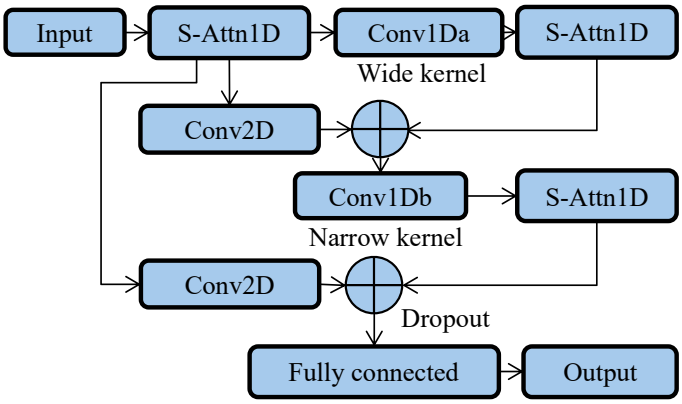

Fig. 3. Hybrid convolutional network architecture employed in this study

To apply bandpass filtering to a discrete signal, normalization of cutoff frequencies is performed relative to the sampling rate Nyquist frequency. This is illustrated below, where fs is the sampling frequency.

$$f_{cutoff\_n} = \frac{f_{cutoff}}{0.5 \times fs} \tag{6}$$

The normalized low f_{low_n} and high cutoffs f_{high_n} are then fed to the Butterworth filter function, and the frequency response is as shown in Fig. 2, where the low cutoff is 8 Hz, and the high cutoff is 32 Hz. The function we applied for this study is as follows:

$$E_{bandpass}(t) = Butter_{scipy}[x(t), f_{low\_n}, f_{high\_n}, f_s, n] \tag{7}$$

where $x(t)$ is the EEG signal at time t, and n is the filter order set to a default value of 4, and $\text{Butter}_{\text{scipy}}$ is the Butterworth bandpass filter designed and applied using the scipy.signal.butter() and scipy.signal.filtfilt() functions from the SciPy signal processing library in Python [27].

3.6. Bandpass filtering with temporal derivative

Bandpass filtering is combined with the 1st-order temporal derivative to achieve the effect of combined denoising. The envelopes are added on top of each other as below:

$$E_{bandpass,derivative}(t) = E_{bandpass}(t) + E_{derivative}(t) \tag{8}$$

3.7. Classification

For classification, we use the lightweight CNN model, originally introduced in our previous work [25]. This architecture is designed for seizure detection and prediction with low computational cost, and it captures both temporal and spatial

Table 1

Impact of preprocessing techniques on LITE model performance

Pre-processing		1 s clip length				5 s clip length			
		Patient		Global		Patient		Global	
		Detect	Detect, Predict	Detect	Detect, Predict	Detect	Detect, Predict	Detect	Detect, Predict
None [25]		0.961	0.834	0.898	0.674	0.812	0.750	0.866	0.813
Ch. sel. [28]		0.965	0.848	0.941	0.731	0.811	0.778	0.960	0.830
Hilbert		not run	not run	0.888	0.683	not run	not run	0.894	0.713
RMS		not run	not run	0.919	0.686	not run	not run	0.911	0.776
Log-power		not run	not run	0.899	0.677	not run	not run	0.901	0.772
Deriv.	f'	0.971	0.886	0.959	0.840	0.713	0.698	0.958	0.867
	f''	0.933	0.859	0.907	0.811	0.744	0.668	0.946	0.827
	f'''	0.928	0.841	0.877	0.762	0.807	0.716	0.914	0.820
BP (Hz)	8 – 32	0.964	0.867	0.947	0.801	0.728	0.739	0.948	0.842
	4 – 13	0.962	0.869	0.938	0.770	0.691	0.690	0.957	0.861
	5 – 25	0.964	0.876	0.946	0.766	0.726	0.717	0.965	0.868
	5 – 35	0.964	0.866	0.941	0.800	0.660	0.696	0.980	0.867
	10 – 25	0.959	0.866	0.921	0.757	0.713	0.712	0.958	0.843
BP (Hz) & f'	8 – 32	0.970	0.881	0.958	**0.851**	0.731	0.733	0.961	0.868
	4 – 13	0.966	0.881	0.947	0.799	**0.800**	0.738	0.958	**0.888**
	5 – 25	**0.971**	0.889	**0.964**	0.832	0.782	**0.783**	**0.982**	0.870
	5 – 35	0.968	0.887	0.956	0.832	0.776	0.708	0.959	0.855
	10 – 25	0.966	**0.890**	0.959	0.841	0.794	0.675	0.976	0.870
Uplift (%)		0.622	4.953	2.444	16.416	-1.356	0.643	2.292	6.988

features by combining a one-dimensional (1D) CNN and two-dimensional (2D) CNN layers that are placed in parallel, as shown in Fig. 3. Therefore, the EEG input is fed into a simple attention 1D block (S-Attn1D) that computes attention weights for each data point in the sequence.

The S-Attn1D consists of two 1D CNN layers, where the first reduces the input channels to one-eighth of the original input size and the second layer uses an activation function to derive attention weights. Finally, the output is then multiplied by the input and fed into a 1D CNN (Conv1Da) without any attention mechanism. This process is repeated twice more in the pipeline, after the Conv1Da and after the Conv1Db, which is derived from the concatenation of a 2D CNN and a 1D CNN.

The model is trained using the cross-entropy loss function. For all classification tasks, 80% of the data is used for training and the remaining 20% for validation. We evaluated model performance on the CHB-MIT dataset for each patient and globally, for all patients, using 1-second and 5-second EEG clips. Two classification tasks are evaluated: 1) seizure detection, where we classify between ictal and non-ictal clips, in total 2 classes, and 2) detection and prediction, where we classify between ictal, preictal, and interictal tasks, in total 3 classes. Although preictal clips are not labeled on the dataset, the smartEEG framework allows us to

label them as preictal when they sit between 1 and 300 seconds before the start of a seizure. The effectiveness of each preprocessing method was assessed using a lightweight CNN classifier, with global and patient-level results reported.

3.8. Experimental Results

Table 1 shows the overall results when applying the preprocessing techniques discussed in this paper. The first two rows indicate results from our previous works [25] and [28], where no preprocessing and the best channel selection techniques were applied. Data for Hilbert, RMS, and Log-power is only available for the global use case, as the results are worse than the baseline and do not constitute a reason to proceed in this direction. Among signal derivatives, the first order derivative f' generally makes the most significant improvement, whereas the 2nd-order derivative f'' and 3rd-order derivative f''' will give out a result below the baseline. Next, the filtering bands generally result in improved performance, especially for 1-second clips and the global class.
Finally, to further boost the performance, both signal filtering of the same bands evaluated and the signal 1st-order derivative f' on top by the addition operation are applied. This returns the best results overall, and the best overall improvement is for the detection and prediction task, for 1-second clips and global use case, followed by the 2nd best improvement for the same task when using 5-second-long clips, globally. The best overall results are recorded for the detection task, globally, 5-second-long clips, with a validation accuracy of 0.982.

Finally, to further boost the performance we apply both signal filtering of the same bands evaluated and the signal 1st-order derivative f' on top by the addition operation. This returns the best results overall, and the best overall improvement is for the detection and prediction task, for 1-second clips and global use case, followed by the 2nd best improvement for the same task when using 5-second-long clips, globally. The best overall results are recorded for the detection task, globally, 5-second-long clips, with a validation accuracy of 0.982.

4. Discussion

Table 2

Effect of clip length and number of samples on overfitting

Patient Config.	Clip length (s)	Detect				Detect and Predict			
		Nr. Clips	Average Accuracy			Nr. Clips	Average Accuracy		
			Train.	Valid.	Δ (%)		Train.	Valid.	Δ (%)
chb15	1	3944	0.995	0.986	-1	5916	0.981	0.877	-11
chb16		148	0.982	0.83	-15	222	1	0.708	-29
Global		22826	0.985	0.956	-3	34239	0.917	0.828	-10
chb15	5	770	0.999	0.985	-1	1155	1	0.874	-13
chb16		22	0.966	0.782	-19	33	0.969	0.66	-32
Global		4398	0.994	0.969	-3	6597	0.988	0.869	-12

Table 3

Average validation accuracy improvement of techniques over channel selection method, 5-second clip long patient-only results are excluded.

Technique	Average Improvement (%)		
	Detection	Detection, Prediction	Overall
No preprocessing	-5.2%	-3.8%	-4.5%
Channel selection	0.0%	0.0%	0.0%
Hilbert envelope	-7.2%	-15.0%	-11.1%
RMS	-4.4%	-9.8%	-7.1%
Log-power	-6.5%	-13.7%	-10.1%
1st-order derivative f'	0.8%	7.1%	3.9%
Bandpass filter 5 to 25 Hz	0.3%	4.0%	2.2%
Bandpass filter 5 to 25 Hz and 1st-order derivative	1.7%	7.0%	4.4%

Further investigation of some of the causes behind obtaining a relatively lower performance with a larger clip length setting for the patient-only use case is made as follows: Patient chb16, due to sparse seizure data, after data balancing, offers the lowest number of clips for any classification task. Severe overfitting on noise, with an ever-increasing validation loss and instability in the model training, is shown in Table 2 and Fig. 4, as opposed to patient chb15 and the global dataset.

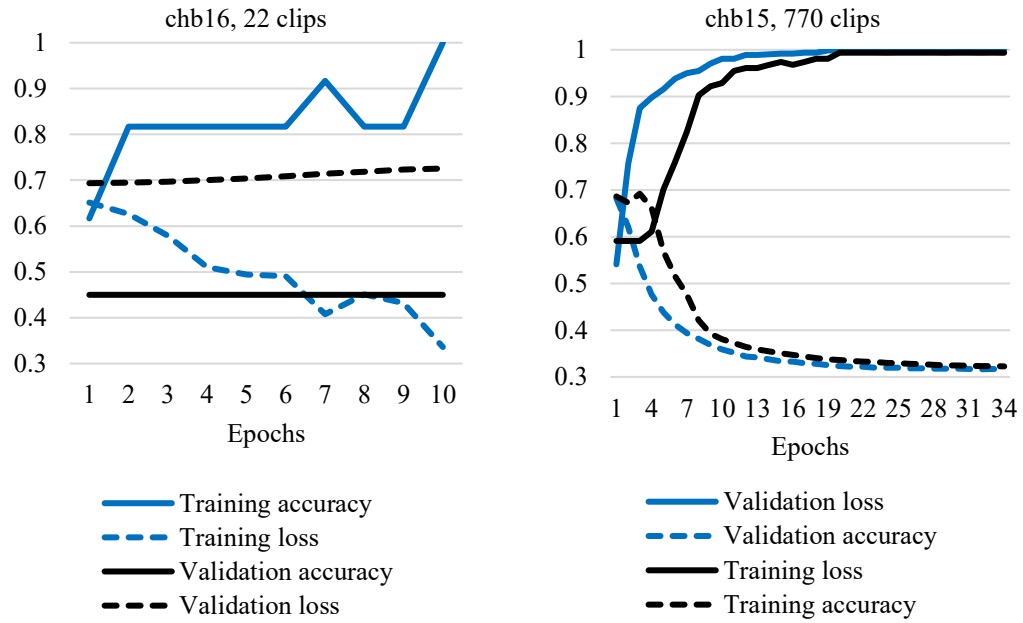

Fig. 4. Comparison of training and validation loss curves for two patients using 5-second EEG clips in the seizure detection task. Left: Patient chb16; Right: patient chb15.

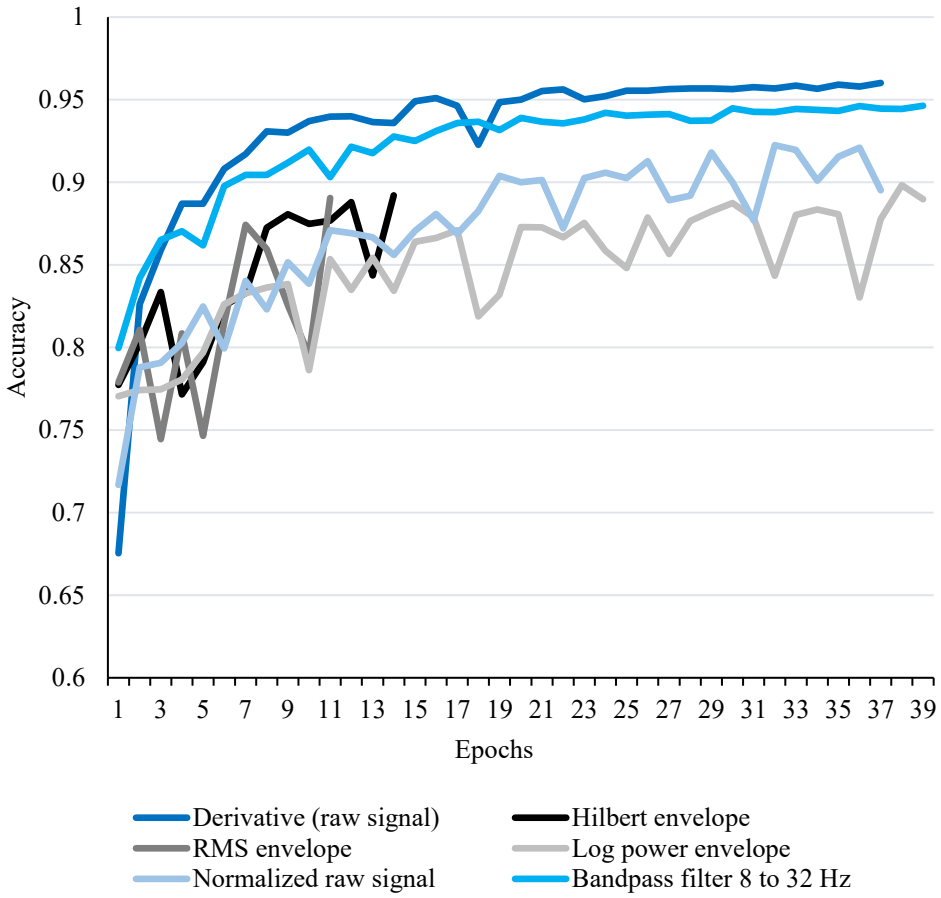

Fig. 5. Comparison of classification validation accuracy across different EEG processing techniques: Hilbert envelope, RMS envelope, log-power envelope, normalized raw EEG signal, and first-order temporal derivative. Each method is evaluated using the global configuration, with 1-second clips in the seizure detection task.

Table 2 compares chb15 and chb16 along with the global use, where the first contains a rich data set, generating the greatest number of clips per patient, the next contains a poor dataset, with the least number of clips per patient, and the final contains a dataset with variable distribution.

The average accuracy is computed for all BP & f' combinations as shown in Table 1. We can clearly see that, on average, training on chb16 will yield severe overtraining, whereas training on a global dataset but with a different distribution, we obtain a result comparable to chb15, which provides a rich dataset within the same distribution.

Table 3 compares improvements between different preprocessing techniques. Training on 5-second clips, on the CHB dataset, provides a lot of

opportunities where training cannot be accomplished due to very low clip counts, thus providing inconsistent results, often prone to severe overtraining, and therefore these results are excluded from Table 3.

Finally, Fig. 5 shows the model training behavior when applying the preprocessing techniques discussed in this paper, showcasing model stability with respect to each technique. The Hilbert Envelope technique transforms EEG raw data into positive-only signals and includes noise information as well. Indeed, in some cases, there's a minor improvement over no preprocessing, but overall, there's an 11.1% decrease in performance compared to channel selection, as showcased by Table 3. The RMS envelope technique provides similar results to no preprocessing and is no longer considered. Finally, the Log Power envelope technique provides similar overall results to the Hilbert envelope. This is most likely occurring due to a significant decrease in signal-to-noise ratio as the log function applies compression. The 1st-order derivative seems to provide a marginal average improvement for detecting seizures, but will allow for discrimination between preictal, non-ictal, and ictal classes with more ease. In principle, the 1st-order derivative will minimize high-frequency noise and accentuate seizure content where there can be high-frequency, large-amplitude changes in the signal, with a benefit of up to 7%. The 2nd-order derivative and 3rd-order derivative were not included in Table 3, as they both provide a lower performance and are not considered. The bandpass filter alone is comparable to Channel Selection, though we can see it's improving more in the three-class task, with benefits up to 4.4% over Channel Selection.

In conclusion, the results suggest a robust and interpretable approach, particularly when the bandpass filtering is combined with the 1st order temporal derivative, as shown in Table 1. This approach generalizes well across both patient-specific and global tasks, without compromising the performance.

Compared to previous works discussed in the Introduction, for the detection task, Chung et al reported a 98.47% accuracy using 4-second clips and applying overlap, while Shen et al. achieved 97.74% accuracy with a duration of 1.35 s and 1s overlap. While these results are high, there's a risk that the validation process introduces redundant samples and leads to data leakage between training and validation sets. Additionally, both papers only use a subset of the entire dataset, with 13 patients for the first and 16 patients out of 24 for the 2nd, respectively. With this in mind, it indicates the authors may not have looked for an algorithm with good generalization that provides some level of immunity to EEG noise and artifacts. Many recent works only look at the seizure detection tasks, adopt long window segments as input, and transform-based preprocessing like DWT, or graph-based architectures, which increase computational complexity and reduce signal interpretability. When using overlapping window data, some of the same EEG waveforms will appear in both the training and the test data. This will result in the

model training on data already available in the test set, thus showing elevated accuracies, but in reality, with a much lower generalization capability.

Our approach avoids these risks by using non-overlapping clips for 1 s and 5 s from all of the subjects in the database, ensuring no leakage information is present and reflecting real-time deployment constraints without compromising the performance, resulting in a pipeline with more generalization capability, accompanied by noise, specific patient distribution, and artifact immunity. A comparison of the best results from patient-specific studies is shown in Table 4.

Table 4

Comparison of CHB-MIT patient-specific EEG seizure classification

Detection Task	# patients	Clip len. [s]	Overlap window	Accuracy
Cao et al. [9]	Not specified			0.984
Shen et al. [12]	16	1.35	Yes (1s)	0.977
Deng et al. [13]	23	0.5 or 1	Yes, step unknown	0.995
Rukhsar et al. [19]	23	0.5 and 1	Yes (25%)	0.963 (0.5 s), 0.959 (1 s)
Chung et al. [20]	13	4	Yes (4s)	0.984
Prediction Task				
Xiang et al. [10]	19	1 and 5	Sliding	0.947 (1 s), 0.982 (5 s)
Our proposal (Detect & Predict)	24 (all)	1	No	0.971 (Det.), 0.890 (Det. & Pred.)

REFERENCES

[1] J. C. Henry, 'Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Fifth Edition', Neurology, vol. 67, no. 11, 2006, doi: 10.1212/01.wnl.0000243257.85592.9a.

[2] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, 'Deep learning-based electroencephalography analysis: A systematic review', 2019. doi: 10.1088/1741-2552/ab260c.

[3] Z. M. Hussain and B. Boashash, 'Hilbert transformer and time delay: statistical comparison in the presence of Gaussian noise', IEEE transactions on signal processing, vol. 50, no. 3, pp. 501–508, 2002.

[4] P. P. Ujma et al., 'The sleep EEG envelope is a novel, neuronal firing-based human biomarker', Sci Rep, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-22255-4.

[5] L. Wang, E. X. Wu, and F. Chen, 'Robust EEG-Based Decoding of Auditory Attention With High-RMS-Level Speech Segments in Noisy Conditions', Front Hum Neurosci, vol. 14, 2020, doi: 10.3389/fnhum.2020.557534.

[6] G. Pfurtscheller and F. H. Lopes Da Silva, 'Event-related EEG/MEG synchronization and desynchronization: Basic principles', 1999. doi: 10.1016/S1388-2457(99)00141-8.

[7] S. Sanei and J. A. Chambers, EEG signal processing. John Wiley & Sons, 2013.

[8] A. H. Shoeb and J. V Guttag, 'Application of machine learning to epileptic seizure detection', in Proceedings of the 27th international conference on machine learning (ICML-10), 2010, pp. 975–982.

[9] X. Cao, S. Zheng, J. Zhang, W. Chen, and G. Du, 'A hybrid CNN-Bi-LSTM model with feature fusion for accurate epilepsy seizure detection', BMC Med Inform Decis Mak, vol. 25, no. 1, p. 6, 2025.

[10] J. Xiang, Y. Li, X. Wu, Y. Dong, X. Wen, and Y. Niu, 'Synchronization-based graph spatio-temporal attention network for seizure prediction', Sci Rep, vol. 15, no. 1, p. 4080, 2025.

[11] A. Ciurea, C.-P. Manoila, and B. Ionescu, 'SmartEEG: An End-to-End Framework for the Analysis and Classification of EEG signals', in 2021 International Conference on e-Health and Bioengineering (EHB), 2021, pp. 1–4.

[12] M. Shen, F. Yang, P. Wen, B. Song, and Y. Li, 'A real-time epilepsy seizure detection approach based on EEG using short-time Fourier transform and Google-Net convolutional neural network', Heliyon, vol. 10, no. 11, 2024.

[13] X. Deng, 'A novel dual-branch network for comprehensive spatiotemporal information integration for EEG-based epileptic seizure detection', PLoS One, vol. 20, no. 6, p. e0321942, 2025.

[14] S. M. Malakouti, M. B. Menhaj, and A. A. Suratgar, 'Enhanced epilepsy detection using discrete wavelet transform and bandpass filtering on EEG data: integration of ART-based and LVQ models', Clinical eHealth, 2025.

[15] 'The Bonn EEG time series download page - Nonlinear Time Series Analysis (UPF)'. Accessed: Jul. 24, 2025. [Online]. Available: https://www.upf.edu/web/ntsa/downloads/-/asset_publisher/xvT6E4pczrBw/content/2001-indications-of-nonlinear-deterministic-and-finite-dimensional-structures-in-time-series-of-brain-electrical-activity-dependence-on-recording-regi

[16] S. A. Karthik, K. N. Bharath, B. R. Ramji, K. Puttegowda, B. Aruna, and D. S. S. Kumar, 'Enhanced EEG Signal Processing for Accurate Epileptic Seizure Detection', SN Comput Sci, vol. 6, no. 6, p. 608, 2025.

[17] V. Shah et al., 'The Temple University Hospital seizure detection corpus', Front Neuroinform, vol. 12, 2018, doi: 10.3389/fninf.2018.00083.

[18] K. Yan et al., 'Automated seizure detection in epilepsy using a novel dynamic temporal-spatial graph attention network', Sci Rep, vol. 15, no. 1, p. 16392, 2025.

[19] S. Rukhsar and A. K. Tiwari, 'Lightweight convolution transformer for cross-patient seizure detection in multi-channel EEG signals', Comput Methods Programs Biomed, vol. 242, 2023, doi: 10.1016/j.cmpb.2023.107856.

[20] Y. G. Chung, A. Cho, H. Kim, and K. J. Kim, 'Single-channel seizure detection with clinical confirmation of seizure locations using CHB-MIT dataset', Front Neurol, vol. 15, p. 1389731, 2024.

[21] A. J. Prakash and M. Atef, 'A lightweight deep learning approach for patient-specific electrocardiogram beat classification using local and long-term dependencies', Eng Appl Artif Intell, vol. 152, p. 110754, 2025.

[22] G. B. Moody and R. G. Mark, 'The impact of the MIT-BIH arrhythmia database', 2001. doi: 10.1109/51.932724.

[23] A. Esmaeilpour, S. S. Tabarestani, and A. Niazi, 'Deep learning-based seizure prediction using EEG signals: A comparative analysis of classification methods on the CHB-MIT dataset', Engineering Reports, vol. 6, no. 11, p. e12918, 2024.

[24] A. Widmann, E. Schröger, and B. Maess, 'Digital filter design for electrophysiological data - a practical approach', J Neurosci Methods, vol. 250, 2015, doi: 10.1016/j.jneumeth.2014.08.002.

[25] A. Ciurea, C.-P. Manoila, and B. Ionescu, 'EEG Classification Using Hybrid Convolutional Neural Network with Attention Mechanism', in International Conference on e-Health and Bioengineering, 2023, pp. 783–791.

[26] K. Majumdar, 'Differential operator in seizure detection', Comput Biol Med, vol. 42, no. 1, 2012, doi: 10.1016/j.compbiomed.2011.10.010.

[27] G. Van Rossum and F. L. Drake, Python 3 Reference Manual. Scotts Valley, CA: CreateSpace, 2009. [28] A. Ciurea, C.-P. Manoila, and B. Ionescu, 'EEG Denoising with Channel Ranking and Selection', in 2024 E-Health and Bioengineering Conference (EHB), 2024, pp. 1–4.