Reducing False Alarms in Seizure Prediction: A Specialized CNN Architecture and a Novel EEG Sampling Technique

Ameer Hayder Drbeeni[1], Haider Ismael Shahadi[2], and Arwa Hameed Mohammed Taqi[3]

[1,2,3]College of Engineering, University of Kerbala, Karbala, Iraq
[1]Email Id: ameer.hayder@s.uokerbala.edu.iq
[2]Email Id: haider_almayaly@uokerbala.edu.iq
[3]Email Id: arwa.h@uokerbala.edu.iq

July 6, 2024

Abstract

Deep learning played a vital role in the seizure prediction challenge. Nevertheless, most studies used generic architectures that fail to consider the distinct characteristics of multivariate time-series Electroencephalography signals. Additionally, many methods depend on inadequate EEG segmentation techniques, resulting in unreliable results. This study presents an in-depth architectural design of a Convolutional Neural Network specifically tailored to extract features from the wavelet-transformed EEG signals using Wavelet packet decomposition (WPD). In addition, the chosen testing strategy and data segmentation methodology ensures accurate and trustworthy performance results. This study introduces a data segmentation method to generate distinct intervals and effectively capture the temporal dynamics of the time-series data. The proposed model evaluation utilized 12 subjects' EEG data from the CHB-MIT dataset, employing a Leave-One-Out cross-validation technique. The proposed architecture outperformed five reproduced state-of-the-art models in the segment-based accuracy, sensitivity, and specificity metrics. The proposed model achieved 78.00% accuracy, 65.17% sensitivity, and a high 90.83% specificity rate. Evaluation using the more straightforward KFold cross-validation technique demonstrated robust performance, achieving 96.68% accuracy, 97.41% sensitivity, and 95.95% specificity. The significant improvement in the model's specificity rates indicates a substantial reduction in false alarms, making the proposed model a reliable tool for seizure prediction.

Keywords: Epilepsy, Seizure Prediction, Electroencephalography, Convolutional Neural Network, Wavelet Packet Decomposition

1 Introduction

Epilepsy is a neurological condition characterized by recurrent and spontaneously occurring seizure events that disrupt and affect patients' lives. The unpredictable nature of seizures can drastically degrade the quality of life and can be life-threatening directly or indirectly owing to drowning and other accidents [1]. The World Health Organization (WHO) reported that epilepsy has affected more than 50 million individuals around the world. The burden of epilepsy is compounded by the fact that 30% of individuals exhibit resistance to anti-epileptic drugs [2]. This burden highlights the imperative need for innovative strategies to predict, detect and prevent seizures [3].

Electroencephalography (EEG) is a common tool for monitoring the brain's electric activity and studying epilepsy [4]. The states of an epileptic brain can be one of four: inter-ictal (normal interval between seizure events), pre-ictal (interval immediately preceding a seizure event), ictal (period of a seizure), and post-ictal (interval immediately following a seizure). At its core, seizure prediction is a task of detecting the presence of the pre-ictal state. However, most studies depict seizure prediction as a binary classification task between the pre- and inter-ictal states [5].

Seizure prediction utilizing the electric activity of an epileptic brain is a challenging research problem that has attracted more attention in recent years [5–8]. The ability to issue warnings before seizures occur may lead to the development of novel diagnostic methods and therapies [9, 10]. Moreover, the focus on the algorithmic prediction of seizures might offer valuable insights into the mechanics of a seizure event. However, little attention has been paid to extracting meaningful and comprehensible characteristics about seizures and their dynamics [11].

EEG signals are complex, non-stationary, and require advanced signal processing and feature extraction to help provide meaningful interpretation [12]. Advances in deep

learning have shown great promise in its capacity to learn good representations of EEG signals. Deep learning approaches differ from traditional machine learning techniques; they do not require extensive domain expertise or operator feedback on the feature extraction process. Nevertheless, deep learning models are prone to weak generalization and deliver less human comprehensible features [13].

Deep learning approaches offer a potential answer to developing a reliable seizure prediction system. However, the clinical adoption of deep learning models mandates a certain degree of model explainability. Although deep learning is considered a black box, it is possible to build human-comprehensible interpretations on the internals of a deep learning model [11, 14]. Nevertheless, most studies utilized deep learning methods with generic architectural designs solely for accurate predictions. Such design choices greatly complicated the interpretation of model decisions. Another concern involving numerous studies is the potential for data bias in the data pre-processing and model evaluation techniques [15].

This study introduces a wavelet-based Convolutional Neural Network (CNN) well-suited for multi-channel EEG data analysis. The architecture's compactness relies on exploiting Wavelet Packet Decomposition (WPD) to enhance the signal for better feature extraction without increasing the data dimensionality or redundancy. The CNN architecture utilizes separable convolutions to reduce the number of learned parameters. Furthermore, the proposed architecture performance was evaluated using the Leave-One-Out Cross-Validation (LOOCV), also named Leave-One-Group-Out (LOGO), for the test set. Additionally, a fixed-step ordered sampling technique was utilized to avoid the uncertainty of random sampling techniques. To ensure a fair comparison, recent state-of-the-art methods are replicated to be trained in the same conditions and to prevent potential data bias.

The remainder of this manuscript is organized as follows: Section 2 gives a concise overview of recent state-of-the-art methods. Section 3 describes the data and materials used in this research, including the employed data pre-processing techniques. Additionally, section 3 lists and details the components of the proposed architecture, and presents the model training strategy and evaluation metrics. Section 4 delivers the performance results, while section 5 thoroughly examines the comparisons of the reproduced methods. Section 6 encloses the manuscript with the study conclusions and discusses potential future works.

2 Related Work

Multiple seizure prediction studies have explored numerous manual feature extraction techniques [16–18]. However, their findings revealed no success in identifying the ideal set of features and the appropriate classifier to yield reliable prediction performance. As a concrete example, the winning study of the Melbourne University seizure prediction contest employed 11 classifiers and more than 3000 handcrafted features [19]. Consequently, recent state-of-the-art methods utilized automated feature extraction, taking advantage of deep learning methods.

Tsiouris et al. [18] used the most common features, including the time domain, frequency domain, correlation, and graph theory features, to generate 643x1 feature vectors for each 5-second long EEG segment. The Long Short-Term Memory (LSTM) network was employed as the classifier for its inherent advantage in processing the sequential EEG time series. They also tested the LSTM network using raw EEG segments as the input in a feature extraction scenario. However, the feature extraction model was unable to achieve comparable results. The primary cause for this low performance observed while using raw EEG data was the model's low capacity and limited computing hardware capability, which prevented any attempt to increase the model size.

Khan et al. [20] proposed a CNN architecture to extract features from the time-frequency domain of the EEG signals. The time-frequency domain was generated using the Continuous Wavelet Transform (CWT) of 1-second raw EEG segments. They argued that using wavelet-transformed EEG segments to include the frequency information achieves deeper feature extraction without additional model capacity. They used the Mexican-hat mother wavelet and 9 dyadic scales to add the geometric frequency range of 0.25 to 128 Hz in the frequency dimension. On the other hand, Truong et al. [21] used the Short Time Fourier Transform (STFT) to transform the 30-second raw EEG segments into the time-frequency domain. The STFT generated segments that encompassed frequencies ranging from 1 to 128 Hz at the cost of losing information in the time dimension. The information loss occurs since STFT uses sliding windows of fixed lengths to generate the amplitude or power spectra.

Daoud and Bayoumi [22] conducted experiments using various deep learning architectures trained on raw EEG data without pre-processing except for the 5-second segmentation process. Their best-performing architecture with the least number of trainable parameters is a Convolutional Auto-encoder (CAE) for feature extraction and a Bidirectional LSTM (Bi-LSTM) for classification. The CAE was trained in an unsupervised manner using data from several subjects to find lower-dimensional representations of the raw EEG data. The Bi-LSTM uses the learned features from the encoder component of the CAE to carry out the ultimate classification task.

Similarly, Xu et al. [23], Jana and Mukherjee [24] introduced end-to-end deep learning models that directly learn from raw EEG data as input. Jana and Mukherjee [24] used 8-second long EEG segments and developed a

CNN architecture with (3×3) convolution kernel size and a combination of (1×3) and (2×3) max-pooling kernel sizes. Different max-pooling kernel sizes are used because the input EEG segments have more temporal than spatial resolution. On the other hand, Xu et al. [23] used 1D convolution kernels with sizes of (1×20) and (1×10) for the first three convolution blocks. Each of these blocks are followed by (1×10) and (1×5) 1D max-pooling operations. The length of these kernel sizes was chosen to capture longer local temporal dependencies with an increased receptive field. The final two convolutional blocks have standard 2D convolution kernels with (3×3) size to capture spatial and temporal characteristics. The EEG segment was set to 20 seconds long in this study.

An alternative method for expanding the receptive field of the convolution process involves using dilated kernels. Hussein et al. [25], Wang et al. [26], Gao et al. [27] incorporated a dilation rate of more than one to enlarge the kernel length while avoiding an increase in the number of weights within a kernel. Hussein et al. [25] divided the EEG data into segments, each having 30 seconds in length. Then, each segment is transformed into time-frequency representations using CWT with 100 scales. However, the wavelet mother and the distribution of the scales were not reported in this study. The illustrated time-frequency images indicated that the utilized scales produce a frequency range of 0.1 to 100 Hz. The suggested model consisted of 30 convolution layers arranged in parallel paths with varying dilation rates and a kernel size of (3×3) or (5×5). While using 30 convolution layers, the network is constructed to be only 3 convolution blocks deep. The suggested CNN architecture was created with numerous learnable parameters, requiring several hours of training for a single patient. Wang et al. [26] used STFT instead of CWT to transform the 30 seconds long EEG segments. Additionally, they used 3D convolution kernels instead of 2D but did not report the number of kernels in each layer.

Lawhern et al. [28] constructed an EEG-specific architecture; called EEGNet, replicating the well-known Filter-Bank Common Spatial Pattern (FBCSP) feature extraction process. FPCSP performs spatial filtering separately on several frequency subbands to compute features with maximal variance to optimally discriminate between two classes [29]. Zhang et al. [30] utilized FBCSP to generate 2D images and achieved state-of-the-art performance using a CNN architecture. In contrast, EEGNet utilized raw EEG data to learn both temporal (frequency-related) and spatial features using a CNN directly instead of FPCSP. The EEGNet architecture incorporates a 1D temporal convolution layer with a kernel size of (1×64), followed by spatial depthwise convolution layers with a kernel size of $(N \times 1)$ where N is the number of channels in the input EEG signal.

However, the EEGNet architecture designed by Lawhern et al. [28] was not specifically trained for a seizure prediction task. Jemal et al. [31] made minor adjustments to the EEGNet architecture and trained the model to predict seizures from 20 seconds long raw EEG segments. The temporal kernel size is (1×128) to extract frequency features starting at 2 Hz from an EEG segment with a 256 Hz sampling rate. The architecture additionally included average-pooling layers with a kernel size of (1×16) following both temporal convolution layers to enlarge the receptive field of the model. However, this study failed to report the number of kernels utilized in each stage of the proposed architecture.

3 Materials and Methods

3.1 Data Acquisition

The proposed CNN architecture is trained using a subset of the CHB-MIT scalp EEG dataset gathered at the Children's Hospital Boston [32, 33]. The dataset comprises 664 recordings, grouped into 23 cases from 22 pediatric patients. Each case is denoted by the code chb_i, where i refers to the case ID. Each case has a summary *.txt* file that specifies the start and end timings, the EEG montage used for each recording, the number of seizure events, and their timing relative to the recording. Each case consists of recordings ranging from 9 to 42. Each recording has 10 minutes up to 4 hours of EEG data duration sampled at 256 Hz. This dataset utilized the international 10-20 EEG electrode placement system with bipolar montage. The proposed method uses only 18 channels common among subjects from the overall 23 EEG channels. The selected EEG channels and their corresponding electrode placements are depicted in Figure 1. Figure 2 shows a 150-minute EEG sample from subject *chb01* preceding the seizure of *chb01_15.edf* file.

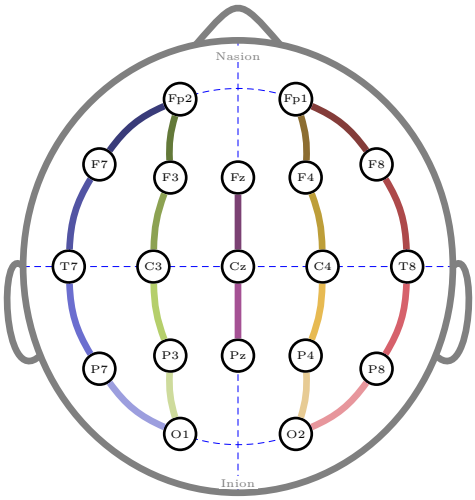


Figure 1: The montage of the 18 common EEG channels among the CHB-MIT subjects.

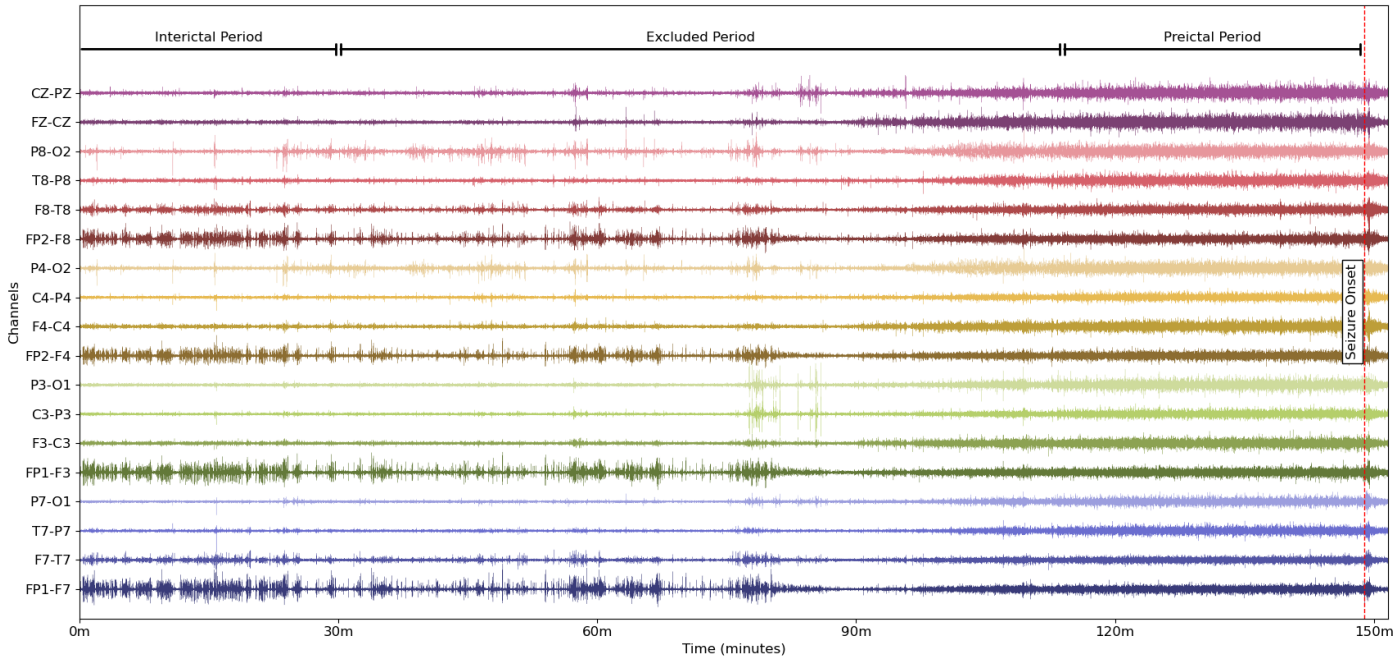


Figure 2: A 150 minutes sample EEG recording from subject *chb01* preceding a seizure.

3.2 Data Segmentation

The proposed seizure prediction system employs the pre- and inter-ictal states to perform a binary classification task. Performing binary classification on long continuous EEG data mandates establishing three parameters to extract the pre- and inter-ictal states:

- The pre-ictal period, which is immediately before a seizure. For this investigation, a 35-minute pre-ictal period length is extracted. Furthermore, seizure events with 30 minutes or less pre-ictal data were not included to ensure non-misleading validation results due to a low number of samples.
- The inter-ictal distance refers to the duration between a seizure onset and the inter-ictal data. This period is not included in the inter-ictal period to reduce the similarity between pre- and inter-ictal states in the training phase. The proposed system uses two hours as the inter-ictal period distance, resulting in an hour and 25 minutes gap between the pre- and inter-ictal periods.
- The post-ictal period is the interval immediately after a seizure that is not categorized as a pre-ictal or inter-ictal state. In this study, two hours of post-ictal period are excluded from the training set.

The intervention time, also known as the Seizure Prediction Horizon (SPH) [10], is another essential parameter required to specify the timings of extracted periods. However, this parameter is only crucial in the event-based prediction (inference stage). For instance, a five-minute intervention time would mean alarms induced in the five minutes immediately before a seizure are false alarms. Nevertheless, the intervention period is still considered pre-ictal; any alarm during this interval is a false warning since there should be enough time to intervene [8, 10]. A 35-minute pre-ictal duration to predict incoming seizures 30 minutes before the intervention interval yields a 30-minute Seizure Occurrence Period (SOP). Only cases with three seizures or more meeting the abovementioned conditions are used to evaluate the proposed and replicated architectures.

Table 1 displays the final subset of EEG data obtained from the CHB-MIT dataset. The table presents the number of seizure events, the overall length of the EEG data, the interictal and preictal data length, and the imbalance ratio between the two classes. The trainable seizures are the number of seizures used in the training process. As shown in Table 1, the inter-ictal data are often long and the pre-ictal data are limited by the pre-ictal length and number of seizure events. The number of inter-ictal samples outnumbering the pre-ictal samples causes a data imbalance issue. To solve the imbalance issue, most studies employed random under-sampling techniques [20, 22–24, 27, 34–36]. Consequently, the certainty of these methods is prone to data bias due to randomly selecting a subset of the inter-ictal data. This study introduce a sampling technique that is formulated as below:

Let the subset of all extracted periods and their annotations from a single case be:

$$S = \{(x_{t,g}, y_{t,g}) | t = 0, 1, ..., N_g - L, g = 0, 1, ..., G\} \quad (1)$$

Table 1: Summary of the CHB-MIT dataset.

Subject	Num. of Seizures (Trainable)	Overall Length (minutes)	Interictal (minutes)	Preictal (minutes)	Imbalance (inter:pre)
chb01	7 (3)	2433	1296	104	12.5
chb02	3 (2)	2115	1564	69	22.7
chb03	7 (1)	2280	1643	33	49.8
chb04	4 (2)	9363	8509	70	121.6
chb05	5 (4)	2340	1233	139	8.9
chb06	10 (6)	4004	2363	209	11.3
chb07	3 (3)	4023	3437	105	32.7
chb08	5 (4)	1200	313	139	2.3
chb09	4 (3)	4072	3360	105	32.0
chb10	7 (5)	3001	1896	175	10.8
chb11	3 (1)	2087	1919	35	54.8
chb12	27 (3)	1241	73	100	0.7
chb13	12 (3)	1980	1202	104	11.6
chb14	8 (4)	1560	584	135	4.3
chb15	20 (8)	2400	428	279	1.5
chb16	10 (2)	1140	458	69	6.6
chb17	3 (2)	1260	907	70	13.0
chb18	6 (3)	2138	1738	104	16.7
chb19	3 (2)	1795	1620	70	23.1
chb20	8 (1)	1656	1202	34	35.4
chb21	4 (2)	1969	1465	68	21.5
chb22	3 (2)	1860	1258	70	18.0
chb23	7 (1)	1593	853	35	24.4

where $x_{t,g}$ is the EEG segment starting at time t and ending at time $t+L$ in group g, L is the selected window size of the EEG segment. $y_{t,g} \in \{0,1\}$ is the binary class label of the corresponding EEG segment. N_g is the total number of EEG time points in group g. G is the total number of groups, where each group represents data associated with a seizure event. While seizure events that met the selection criteria have pre-ictal data of no less than 30 minutes, the seizure events is not necessarily preceded by inter-ictal data. Therefore, the inter-ictal recordings are divided evenly among the seizure events. In this study, the segmentation window size (L) was set 4 seconds. Now let the majority class (inter-ictal class) and the minority class (pre-ictal class) be represented as follows:

$$\begin{aligned} S_{0,g} &= \{(x_{t,g}, y_{t,g}) | y_{t,g} = 0\} \\ S_{1,g} &= \{(x_{t,g}, y_{t,g}) | y_{t,g} = 1\} \end{aligned} \tag{2}$$

where $S_{0,g}$ is the inter-ictal set and $S_{1,g}$ is the pre-ictal set. The complete set S represents all unique segments by hoping only one sample at a time. Tsiouris employed a variable overlapping factor using variable hopping size per subject. Using a large overlap percentage increases the number of samples in the minority class. However, it does not add new information to the dataset and increases the computational cost of the training. Therefore, This study used a fixed overlap percentage of 50%.

$$S_{1,g}^{'} = \{(x_{t,g}, y_{t,g}) | y_{t,g} = 1, t = 0, n, 2n, ..., N_g - L\} \tag{3}$$

where $n = 0.5 * L$ is an integer defining the hop size and equals half the segment size. In order to under-sample the majority class, we used a variable underlapping factor. In contrast to random sampling, this strategy ensures that the chosen subset of the majority class contains as much temporal information as possible. The underlapping hop size calculation is as follows:

$$k = \lfloor \frac{S_{0,g}}{S_{1,g}^{'}} \rfloor \tag{4}$$

Where $\lfloor \cdot \rfloor$ is the floor function. The final under-sampled majority class is:

$$S_{0,g}^{'} = \{(x_{t,g}, y_{t,g}) | y_{t,g} = 0, t = 0, k, 2k, ..., N_g - L\} \tag{5}$$

3.3 Wavelet Packet Decomposition

Wavelet-based and other frequency-time transformations can effectively analyze non-stationary signals like EEG [17, 37]. Both Wavelet Packet Decomposition (WPD) and Discrete Wavelet Transform (DWT) decompose the input

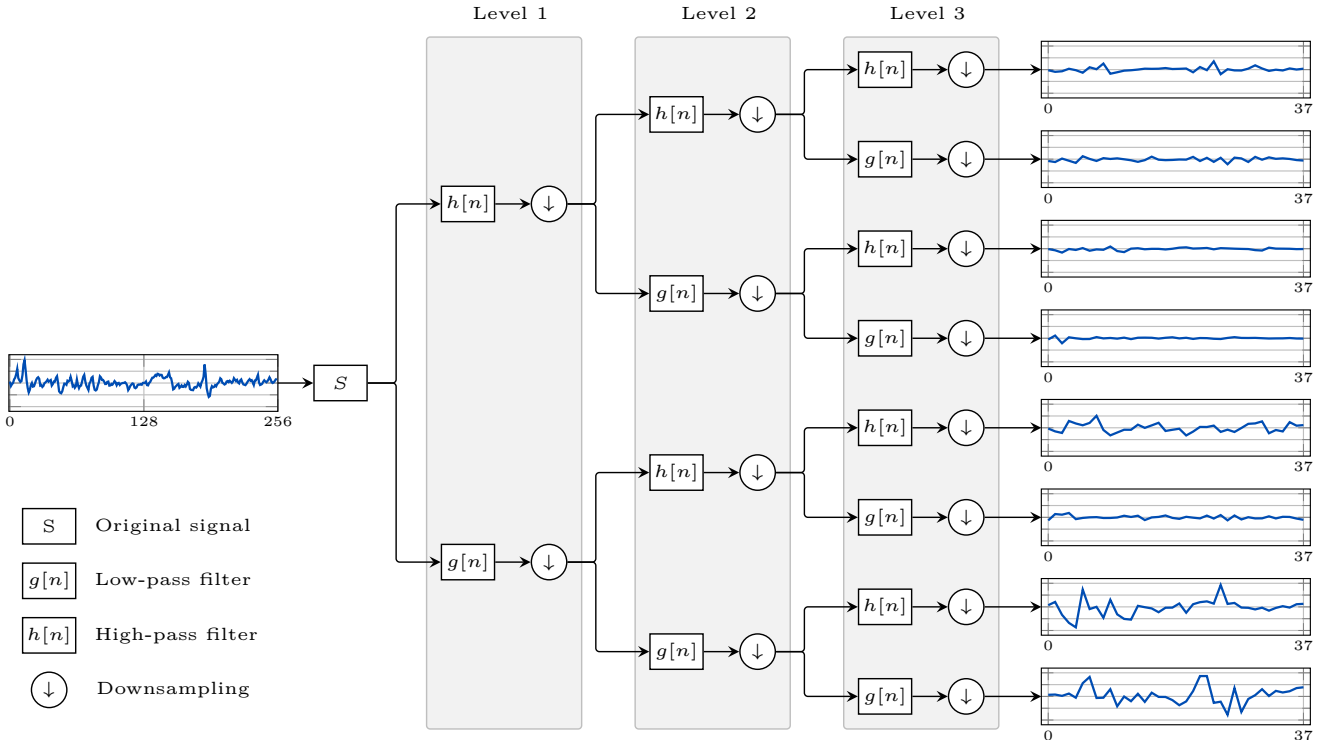


Figure 3: An example of 3-level WPD tree on a single channel of an EEG segment.

signal into a pair of coefficients that reflect the low (approximation) and high-frequency (detail) components of the signal [38]. However, the distinction is that DWT uses only the previous approximation to generate the next-level approximation and detail coefficients. In contrast, WPD decomposes the previous level's approximation and detail coefficients into four new coefficients. Hence, DWT generates $(l+1)$ coefficients whereas WPD produces (2^l) coefficients.

DWT and WPD transformations do not add to the redundancy of the input signal due to the down-sampling process at each level. However, WPD coefficients have equal lengths due to undergoing the same number of down-sampling operations. Conversely, DWT has different time points between the approximation and detail coefficients because of the multi-level process. Hence, the equal-length coefficients generated by WPD are more appropriate for input to a CNN architecture. Figure 3 illustrates a 3-level WPD tree on a single channel of an EEG segment.

Let the chosen segment length be $L = 4$ seconds, where the number of time points in the segment would be $L \times f_s = 4 \times 256 = 1024$ time points. Applying a 4-level WPD using *Daubechies 4* (db4) wavelet to the segment gives:

$$wpd : \mathbb{R}^{18\times1024} \rightarrow \mathbb{R}^{16\times18\times70}$$

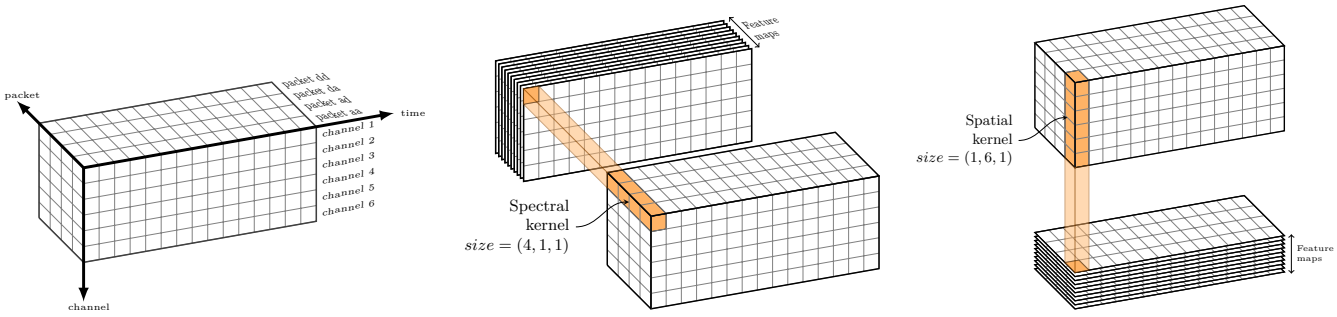
where the number of generated coefficients for each channel is $2^4 = 16$. The down-sampling process at each level reduces the number of time points to 70. Note that $1024/2^4 = 64 \neq 70$ because the db4 filter has a length of 8 points, causing the addition of 7 points to each level in the filtering process. Therefore, the down-sampling process of 4-level WPD yields the following sequence of coefficient lengths for each level:

$$1024 \xrightarrow{L1} \frac{1024+7}{2} = 515 \xrightarrow{L2} 261 \xrightarrow{L3} 134 \xrightarrow{L4} 70$$

3.4 Proposed CNN Architecture

Recently, one-dimensional (1D) CNN architectures have demonstrated exceptional performance in various signal processing applications, including the analysis of univariate signals such as ECG [39] or multivariate data such as EEG [28]. A significant advantage is the simplified interpretation of features extracted using 1D kernels on time-series data. For instance, employing a 2D kernel in an image classification task can be visually explained using basic geometry concepts such as edges and diagonal lines [40].

Similarly, 1D kernels can be analyzed in terms of temporal patterns or frequency information when utilized to extract features from a time series. Additionally, a 1D convolution along the channel axis of a multi-channel signal can indicate the channel that contributes the most to the extracted feature map. Hence, the developed CNN architecture comprises only 1D kernels using separable and depthwise convolutions. The proposed CNN architecture accepts a 3D matrix as input of the shape $(D \times C \times T)$, where the first dimension (D) represents the stacking of



(a) A training sample (b) Spectral Convolution (c) Spatial Convolution

Figure 4: An example of 6-channel 4-coefficient sample to illustrate the process of: (a) The structure of a single training sample, (b) The spectral convolution block, where the convolving kernel slids along the channel and temporal dimenstion to acount for the same wavelet coefficients across individual channels. (c) The spatial convolution block, where the kernel slids over individual chaanels. The packet (aa), (ad), (da), and (dd) represent the equally divided low to high frequency bands of 4-level WPD, respectively.

wavelet coefficients, (C) is the stacked EEG channels, and T is the number of time points in the coefficients. Figure 4a shows an example of a 6-channel 4-coefficient sample.

3.4.1 Spectral Convolution Branch

The first convolution block, termed *Spectral Convolution Block*, applies a 1D kernel along the wavelet nodes of each channel separately, without interference between different channels. Each kernel generates different weighted combinations of the wavelet approximation and details coefficients. The spectral convolution process has a kernel size of $(D \times 1 \times 1)$ with no padding. Therefore, the output of the spectral convolution block has the shape of $(C \times T \times F)$, where F is the number of feature maps (number of filters). Figure 4b illustrates the process of the spectral convolution block.

The second layer, named *Tempral Covolution*, performs a separable convolution on the generated feature maps to extract deeper frequency-related information. The separable convolution comprises a 1D kernel that is one second long to capture frequency information at 1 Hz and above. A level 5 WPD generates an approximation coefficient with a bandwidth of 1 Hz up to $128/2^4 = 4$ Hz. However, the first convolution block aggregate wavelet details with different band-pass frequencies to include the frequency information from the overall 1-128 Hz bandwidth. In this sense, the temporal kernel receptive field spans a duration of 1 second and can extract frequency information of the range (1-128 Hz) while consisting of only (16) samples. The second part of the temporal convolution performs a pointwise convolution to aggregate the extracted temporal features from different wavelet coefficient combinations and generate the final feature maps.

Until now, the network has not included any convolution across individual channels to extract the spatial information. The depthwise convolution block applies a 1D kernel along the channel dimension C left after the spectral convolution block. In the depthwise convolution, the network uses a 1D kernel of size $(C, 1)$ that only slides along the temporal dimension T. The depthwise layer learns a frequency-specific spatial filter for each feature map produced by the spectral convolution block. The depthwise convolution block has a parameter called *depth multiplier*. The depth multiplier controls the number of output channels that the depthwise convolution produces. We set the depth multiplier to 4 to learn four spatial filters for each feature map separately. An average pooling layer of kernel size $(1 \times 1 \times 2)$ follows the depthwise convolution to reduce the temporal dimension further.

3.4.2 Spatial Convolution Branch

The proposed architecture has a second branch that starts with a channel-wise convolution. The channel-wise convolution block, named *Spatial Convolution Block*, performs a 1D convolution operation across EEG channels, ensuring no interference between frequency bands. The prevention of interference between wavelet coefficients is obtained by applying a 1D kernel with a $(1, C, 1)$. The spatial convolution block learns a spatial filter across different frequency bands to generate feature maps of shape $(D \times T \times F)$. The spatial convolution branch consists of layers similar to the spectral convolution branch layers, except that the depthwise convolution has a kernel size of $(D, 1)$ in the spatial branch. The second branch is concatenated with the first branch to form the final feature map. Figure 4c explains the process of the spatial convolution block.

Ultimately, the network applies a 1D temporal convolution to allow the accumulation of concepts learned across the temporal dimension with a kernel half the size of the previous temporal kernel. The final feature map then

passes through two dense layers with units equal to 32 and 1, respectively. The network's output is derived using a *sigmoid* activation function, resulting in a single scalar value representing the predicted class probability. Figure 5 presents complete configuration details of utilized layers of the proposed CNN architecture developed in this study.

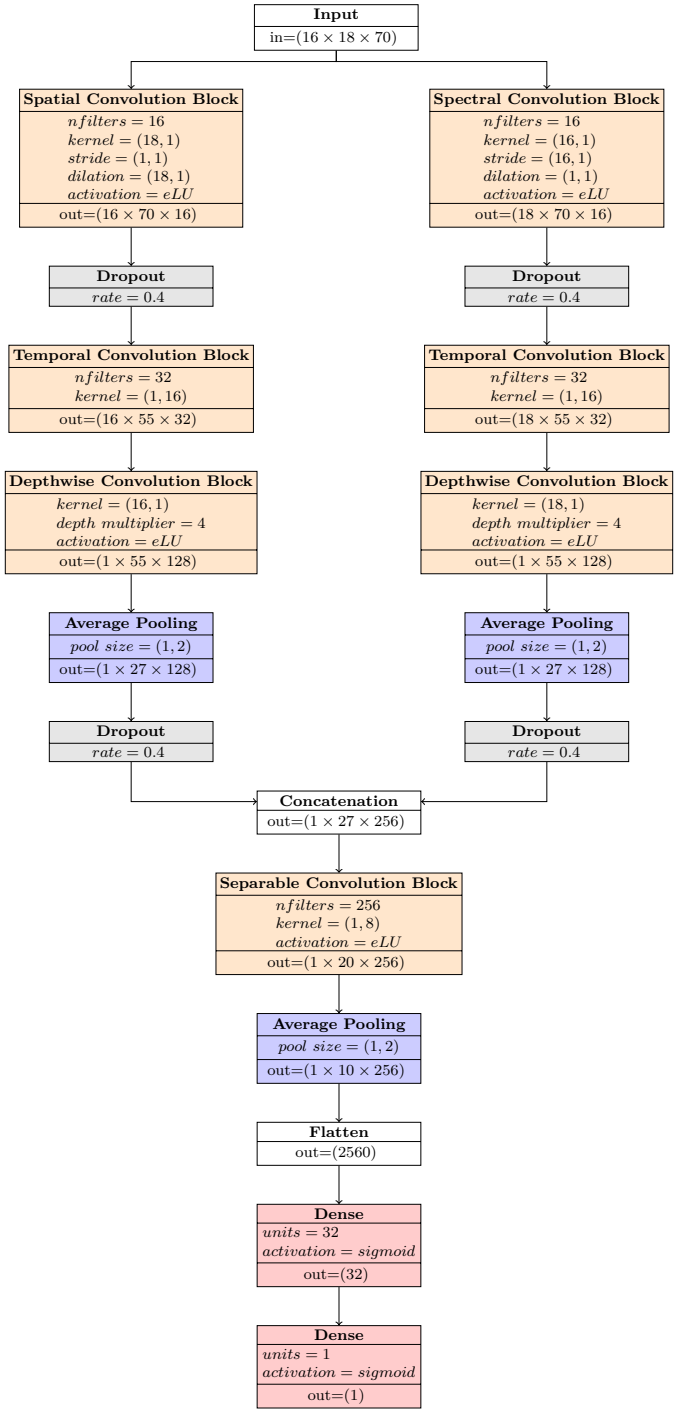


Figure 5: The proposed CNN architecture.

3.5 Training Strategy

The proposed CNN architecture is trained in a subject-specific scheme. The EEG data of a subject is divided into groups, as discussed in section 3.2, where each group involves the data related to a seizure event. The designed CNN architecture was trained and tested using the Leave-One-Group-Out (LOGO) cross-validation method. In this method, the EEG data of one group is used as a test set, leaving the rest of the groups as a training set. If the subject has N seizure events, N models are trained using $(N-1)$ seizure events data and tested on the remained seizure event data. Ultimately, the LOGO cross-validation for the test data split provides robust evaluations of the model generalization to unseen data. Figure 6 illustrates the LOGO test data split method. On the contrary, the K-Fold cross-validation method split the data into K sets randomly, which causes the model to be evaluated on mixture of data from different seizure events.

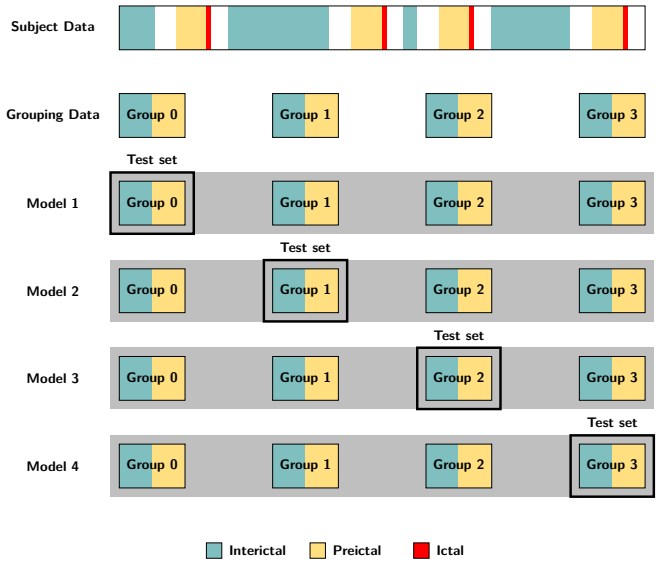


Figure 6: Grouping and test data split using the Leave-One-Group-Out cross-validation method.

The training set is further divided into training and validation sets. While most studies employed a randomly chosen holdout validation set, we used an ordered selection of the validation set. The validation set uses every fifth segment within a class from the training set to construct the 20% validation set. Compared to the random selection of validation samples, the proposed split method ensures that the training and validation sets hold samples with similar temporal characteristics. Additionally, the ordered selection delivers deterministic results, and the training process becomes repeatable.

The holdout validation set prevents the model from over-fitting using an early stopping condition. The parameter *validation patience* is configured to 25, which halts the

training if the validation loss shows no improvement after 25 epochs. The model training utilizes an Adam optimizer with a 0.001 learning rate and a 512 batch size. The model is trained for 500 epochs or until violating the early stopping criteria. Additionally, the model with the best parameters before halting the training process was restored to provide the evaluations on the test set. Given that data labels are binary, model training employs the binary cross-entropy loss function. This study discusses only the segment-based results, as the event-based results can be easily derived by applying any post-processing method. The accuracy, sensitivity, and specificity metrics were used to assess the model's performance compared to the replicated models.

4 Results

The designed CNN architecture and the developed data segmentation method was tested using EEG data obtained from 67 seizure events from 12 patients subset of the CHB-MIT dataset. The proposed architecture was compared with five models replicated from the studies of Khan et al. [20], Truong et al. [21], Xu et al. [23], Lawhern et al. [28], Jemal et al. [31]. We closely replicated the same architecture details and hyperparameters as the original studies. Nevertheless, there are still some differences that were not taken into account. For instance, Khan et al. [20] employed the ictal class to train the model in a 3-class classification task. The replication process was conducted to deliver a comparison using the same data segmentation and cross-validation techniques presented in this study. Using the same cross-validation method ensures that the results are comparable and that the data selection does not bias the model performance. Table 2 presents the comparison results of the proposed architecture and the replicated models. The table presents the number of epochs, the elapsed time, and the performance metrics of the models using KFold and LOGO cross-validation techniques.

The fastest converging method was the CNN model of Truong et al. [21], where the proposed architecture is 3 convolution layer deep. Our proposed CNN architecture has the second fastest convergence training time. The table also presents the performance metrics in two cross-validation methods: Random data splitting using K-Fold cross-validation and the robust LOGO cross-validation methods. In the K-Fold cross-validation, the proposed architecture achieved an accuracy, sensitivity and specificity of 96.68%, 97.41% and 95.95%, respectively. However, the Randomness of the data splitting may introduce bias in the model performance and may not provide deterministic results. The replicated methods dropped down to around 60% accuracy in LOGO cross-validation. The largest drop in performance was observed in the sensitivity metric, where the seizure events are from intervals with distinct temporal characteristics. The proposed architecture provided significant improvement over other models in the LOGO cross-validation method. Our architecture achieved an accuracy of 78.00%, sensitivity of 65.17%, and specificity of 90.83%. The second highest sensitivity was achieved by Khan et al. [20] CNN model with 47.97%.

5 Discussion

Sensitivity quantifies the model's potential to detect a pre-ictal segment accurately, whereas specificity measures the model's accuracy in classifying inter-ictal segments. Therefore, it is crucial to enhance sensitivity to improve the ability to predict a seizure occurrence. However, enhancing the specificity is essential to avoid inaccurate pre-ictal classifications, which might lead to false alarms. The model successfully predicted all seizure occurrences in LOGO cross-validation in subjects *chb01, chb06, chb08, chb12 and chb14*, with a sensitivity of $94\% \pm 1.5\%$. Nevertheless, the improvement in model performance compared to the replicated methods can be attributed to the fact that the model is reaching a specificity of $90.83\% \pm 1.6\%$ besides yielding comparable or higher sensitivity rates. Except for subject *chb14*, which has the lowest specificity of $73.89 \pm 5.6\%$, the proposed architecture consistently reached a specificity of $94.13\% \pm 0.9\%$. The only downside of this architecture requiring further investigation is subject *chb18*, which has a 0% sensitivity. Figure 7 showcases the subject-specific performance metrics obtained using the LOGO cross-validation technique to examine the performance of the proposed architecture.

5.1 architectural design choices

The replicated methods were selected based on the availability of the architecture design details and the reproducibility of the pre-processing technique. These studies also employed various pre-processing techniques to compare the performance of WPD with different transformation methods. For instance, Khan et al. [20] utilized CWT to extract the frequency-related features, whereas Truong et al. [21] used STFT. The CNN architectures of these studies employed the EEG channel dimension as the input feature map channel dimension. Hence, the first layer's kernel of these models must have as many channels as the EEG input segment, resulting in a wide spatial receptive field. The kernel sizes utilized were either (3×3) or (5×5), resulting in a low receptive field in the frequency and temporal dimensions. Truong et al. [21] used an FFT window hop size of around 133 samples to down-sample $(30 \times 256 = 7680)$ time points in the original 30-second EEG segment to only 59 time points in the temporal dimension. The down-sampling of STFT leads to an expansion of the temporal receptive field of the model. However, this trade-off causes a substantial loss of temporal information to accommodate the frequency

Table 2: The results of the proposed method compared to the replicated architectures.

Author	Avg. training time (seconds)	KFold CV			LOGO CV		
		Acc	Sen	Spec	Acc	Sen	Spec
Khan et al. [20]	243	80.02	82.00	78.05	60.80	47.97	73.62
Truong et al. [21]	**229**	92.91	93.84	91.98	59.98	39.71	80.25
Lawhern et al. [28]	487	86.94	85.93	87.95	65.24	47.85	82.62
Xu et al. [23]	470	93.06	96.49	89.64	59.62	37.97	81.28
Jemal et al. [31]	332	90.38	90.71	90.04	61.64	41.95	81.34
This study	233	**96.68**	**97.41**	**95.95**	**78.00**	**65.17**	**90.83**

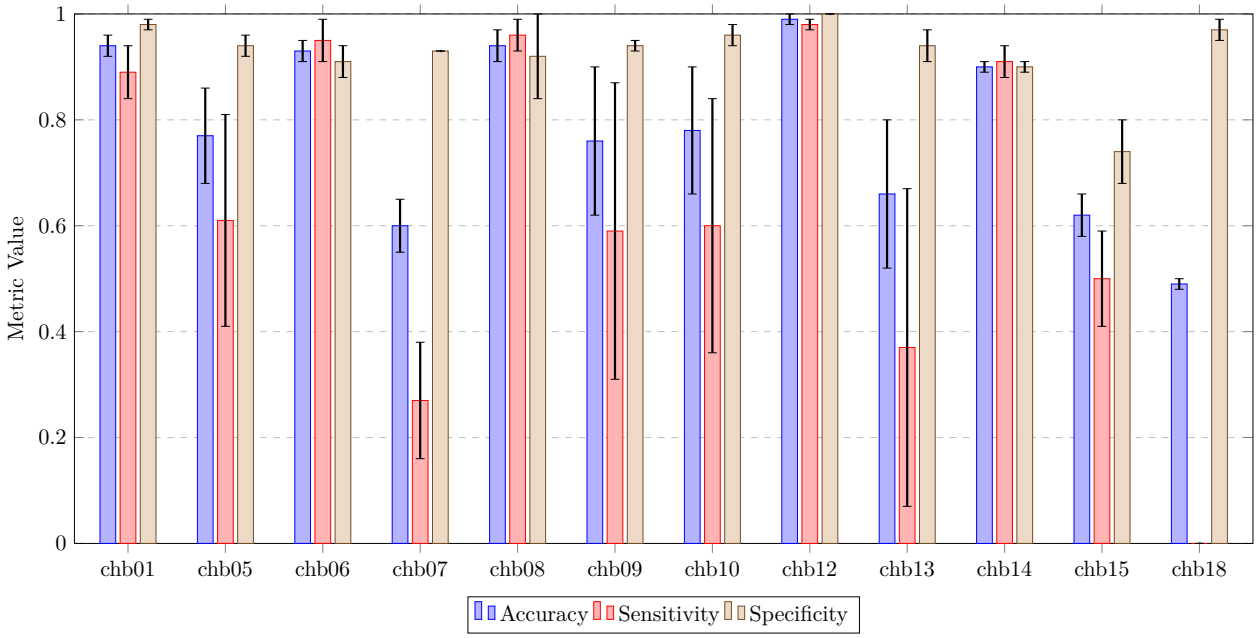


Figure 7: The subject-specific accuracy, sensitivity and specificity metrics of the proposed architecture using LOGO cross-validation. The error bars represent the standard error of the performance metrics.

dimension.

Xu et al. [23] utilized unprocessed EEG segments directly as the input for their model. The model architecture utilized lengthy 1D max-pooling kernels in the last convolution layer, with a cumulative temporal divisor of 1000. Using large max-pooling kernel sizes, specifically (1×10), expanded the temporal receptive field. However, this came at the expense of sacrificing the temporal resolution. The model underwent training using EEG segments of 20 seconds length. Therefore, the last convolution layer achieved a temporal receptive field of 12 seconds, using a kernel size of (3×3). However, the spatial receptive field was not adequately developed since there was no pooling in the channel dimension, except in the last two convolution layers. Jemal et al. [31] utilized a max-pooling kernel size of (1×16) after the first two convolution layers. The architecture of EEGNet utilized a depthwise convolution kernel of (18×1) kernel size to extract the spatial features. Hence, the spatial receptive field equals the number of channels in the EEG segment.

Our model employs the wavelet coefficients of WPD to further extract the frequency-related information. Furthermore, WPD down-samples the time domain of the EEG segment without losing any temporal information. This concept is embedded in the fact that WPD is lossless and can fully reconstruct the original signal from the coefficients. Hence, WPD enhances the representation of the EEG segments, isolates the frequency subbands in the segments, and reduces the temporal dimension. To address the limitations of replicated CNN designs, the proposed CNN architecture has 1D kernels with large receptive fields along the temporal, frequency subbands and channel dimension.

The proposed architecture consists of two branches that carry out distinct aggregation of features. The first branch extracts temporal features aggregated from enhanced coefficients across the channels. In contrast, the second branch extracts temporal features aggregated from enhanced channels utilizing the coefficients. The third layer of the architecture uses a depthwise kernel to further extract the spatial features in the first branch and coefficient features

from the second branch. Thus, this design allows the model to have a receptive field equal to the number of coefficients and channels. Additionally, the down-sampling process of WPD enables the model to have a larger temporal receptive field without increasing the kernel size or reducing the temporal resolution of the input segment. Table 3 presents the model architecture and the employed input nature of the replicated methods and the proposed method. The table also presents the number of parameters in the CNN model, the preprocessing method, the segmentation window size, and the input shape of the EEG segment.

5.2 Performance comparison

The developed model achieved slightly higher or comparable performance for most subjects. Nevertheless, the proposed CNN model achieved significantly higher AUC values for subjects *chb05, chb06, chb07, chb09, chb14, and chb15*. Additionally, we can observe that for subject *chb01, chb13, chb14*, only the original EEGNet could deliver comparable results to the proposed model. The only limitation of the proposed architecture is the inability to classify any pre-ictal segments of subject *chb18*. This limitation could be due to the large gap between the recordings (around 50 hours of missing EEG data from the 90 hours of the subject session). Figure 8 presents the AUC metric values of the proposed architecture and the replicated models using the LOGO cross-validation technique.

The seizure prediction task is highly challenging due to the high similarity between the pre- and inter-ictal segments. Additionally, it has been observed that the longer the temporal distance between the training and testing data result in a decrease in the model's performance. Therefore, the accuracy drop in the specificity is lower than that of the sensitivity. Using an underlapping factor to under-sample the longer inter-ictal data segments improved the model's performance by including more temporal characteristics. However, this holds true only for the inter-ictal class where there are much longer periods available in the subject's data. Therefore, to increase the seizure prediction performance, the model should be trained on longer periods of pre-ictal data.

6 Conclusion and Future Work

This study presented a novel wavelet-based 1D CNN architecture to develop a segment-based seizure prediction model. Considering the nature of EEG signals, The proposed model employed Wavelet Packet Decomposition to enhance the representation of the EEG signals. Additionally, WPD can extract frequency-related details without increasing the complexity or redundancy of the EEG data. Moreover, the WPD pre-processing technique can reduce the temporal dimension in a lossless manner. Therefore, it is possible to enlarge the receptive field of the model of small-sized kernels without sacrificing the resolution when using traditional down-sampling techniques such as pooling or kernel dilation. The proposed model utilized separable and depthwise convolution layers to further enhance and extract deeper features from the wavelet packets.

For a fair comparison with state-of-the-art CNN-based methods, we replicated five models from recent studies. Analysis of the replicated models' architectural limitations motivated the development of the proposed model. The proposed design outperformed the replicated models, significantly improving the seizure prediction ability. The models were trained using two cross-validation methods: Random K-fold and Leave-One-Group-Out. While K-Fold delivers nondeterministic results, the LOGO cross-validation technique showcases a practical testing situation and assesses the model generalization ability. the CHB-MIT dataset is utilized to benchmark the proposed model and the replicated methods. The proposed model outperformed the replicated models, proving the robustness and reliability of the proposed architecture.

The EEG data typically contains long intervals of inter-ictal data, while seizure events with valid pre-ictal periods are rare, leading to a high data imbalance. Most studies utilized random under-sampling techniques to reduce the quantity of inter-ictal samples. This study introduced an underlapping factor to under-sample the inter-ictal data samples while maximizing the incorporation of temporal characteristics. The suggested under-sampling method improved the certainty of the model's performance and increased the specificity of the model, resulting in infrequent false alarms. However, this under-sampling technique only applies to the inter-ictal class, where the data is abundant. To further improve the model's performance and sensitivity, the model's training requires more pre-ictal periods.

A medical application of seizure prediction models requires a high level of explainability and interpretability. While the proposed architecture has an explainable design, further investigation using explainable AI techniques is required to understand the model's decision-making process. Further design enhancements require employing the attention mechanism to focus on the most informative parts of the wavelet-transformed EEG signals. The attention module can also deliver a degree of explainability to the model's predictions.

References

[1] Orrin Devinsky, Annamaria Vezzani, Terence J. O'Brien, Nathalie Jette, Ingrid E. Scheffer, Marco De Curtis, and Piero Perucca. Epilepsy. *Nature Reviews Disease Primers*, 4(1):18024, May 2018. ISSN

Table 3: Summary of the proposed methods and the replicated arhcitectures.

Author	Num. of Params.	Preprocessing	Segmentation	Input shape
Khan et al. [20]	188K	CWT	1s	$10 \times 128 \times 18$
Truong et al. [21]	195K	STFT	30s	$115 \times 59 \times 18$
Lawhern et al. [28]	2.1K	Raw	1s	18×256
Xu et al. [23]	942K	Raw	20s	18×5120
Jemal et al. [31]	2.8K	Raw + filtering	5s	18×1280
This study	160K	Raw	4s	288×70

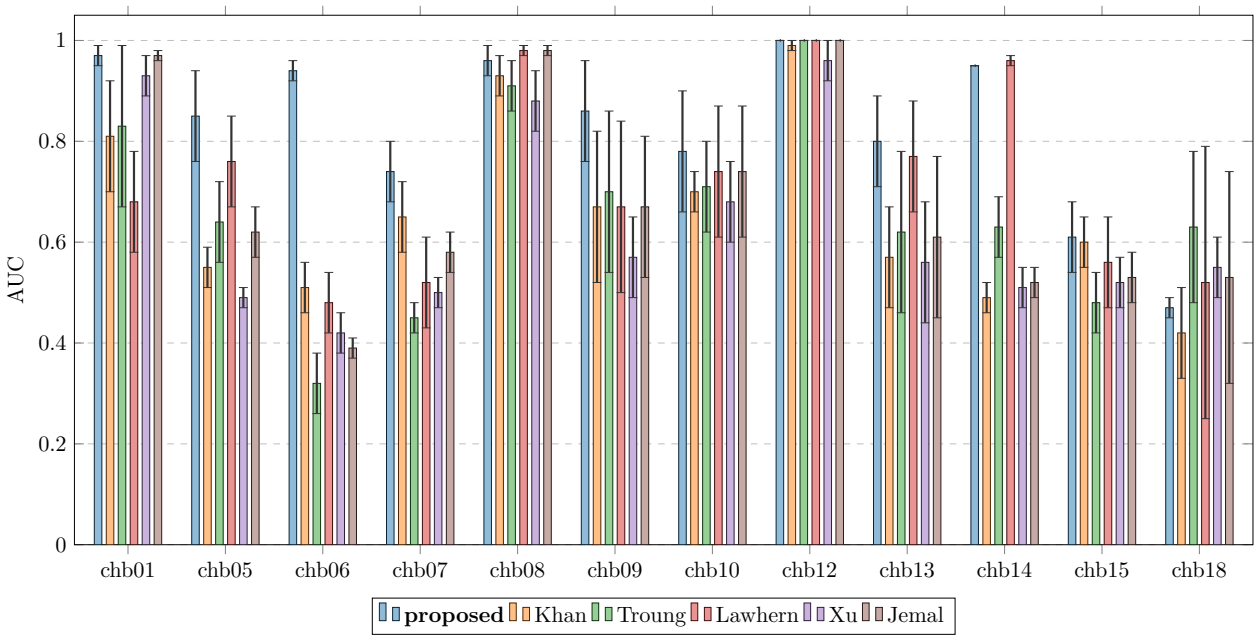


Figure 8: The subject-specific AUC values of the proposed architecture and the replicated methods using LOGO cross-validation. The error bars represent the standard error of the AUC values.

2056-676X. doi: 10.1038/nrdp.2018.24.

[2] World Health Organization. *Epilepsy: A Public Health Imperative*. World Health Organization, Geneva, 2019. ISBN 978-92-4-151593-1.

[3] William C Stacey and Brian Litt. Technology Insight: Neuroengineering and epilepsy—designing devices for seizure control. *Nature Clinical Practice Neurology*, 4 (4):190–201, April 2008. ISSN 1745-834X, 1745-8358. doi: 10.1038/ncpneuro0750.

[4] W.O. Tatum, G. Rubboli, P.W. Kaplan, S.M. Mirsatari, K. Radhakrishnan, D. Gloss, L.O. Caboclo, F.W. Drislane, M. Koutroumanidis, D.L. Schomer, D. Kasteleijn-Nolst Trenite, Mark Cook, and S. Beniczky. Clinical utility of EEG in diagnosing and monitoring epilepsy in adults. *Clinical Neurophysiology*, 129(5):1056–1082, May 2018. ISSN 13882457. doi: 10.1016/j.clinph.2018.01.019.

[5] Khansa Rasheed, Adnan Qayyum, Junaid Qadir, Shobi Sivathamboo, Patrick Kwan, Levin Kuhlmann, Terence O'Brien, and Adeel Razi. Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review. *IEEE Reviews in Biomedical Engineering*, 14:139–155, 2021. ISSN 1937-3333, 1941-1189. doi: 10.1109/RBME.2020.3008792.

[6] Sriram Ramgopal, Sigride Thome-Souza, Michele Jackson, Navah Ester Kadish, Iván Sánchez Fernández, Jacquelyn Klehm, William Bosl, Claus Reinsberger, Steven Schachter, and Tobias Loddenkemper. Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy. *Epilepsy & Behavior*, 37:291–307, August 2014. ISSN 15255050. doi: 10.1016/j.yebeh.2014.06.023.

[7] Kais Gadhoumi, Jean-Marc Lina, Florian Mormann, and Jean Gotman. Seizure prediction for therapeutic devices: A review. *Journal of Neuroscience Methods*, 260:270–282, February 2016. ISSN 01650270. doi: 10.1016/j.jneumeth.2015.06.010.

[8] Levin Kuhlmann, Klaus Lehnertz, Mark P. Richardson, Björn Schelter, and Hitten P. Zaveri. Seizure prediction — ready for a new era. *Nature Reviews Neurology*, 14

(10):618–630, October 2018. ISSN 1759-4766. doi: 10.1038/s41582-018-0055-2.

[9] Leon D. Iasemidis. Seizure Prediction and its Applications. *Neurosurgery Clinics of North America*, 22(4):489–506, October 2011. ISSN 10423680. doi: 10.1016/j.nec.2011.07.004.

[10] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz. Seizure prediction: The long and winding road. *Brain*, 130(2):314–333, February 2007. ISSN 0006-8950, 1460-2156. doi: 10.1093/brain/awl241.

[11] Mauro. F. Pinto, Adriana Leal, Fábio Lopes, José Pais, António Dourado, Francisco Sales, Pedro Martins, and César A. Teixeira. On the clinical acceptance of black-box systems for EEG seizure prediction. *Epilepsia Open*, 7(2):247–259, June 2022. ISSN 2470-9239, 2470-9239. doi: 10.1002/epi4.12597.

[12] Saeid Sanei and J.A. Chambers. *EEG Signal Processing*. Wiley, 1 edition, September 2007. ISBN 978-0-470-02581-9 978-0-470-51192-3. doi: 10.1002/9780470511923.

[13] Zhe Ren, Xiong Han, and Bin Wang. The performance evaluation of the state-of-the-art EEG-based seizure prediction models. *Frontiers in Neurology*, 13:1016224, November 2022. ISSN 1664-2295. doi: 10.3389/fneur.2022.1016224.

[14] Arun Rai. Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, 48 (1):137–141, January 2020. ISSN 0092-0703, 1552-7824. doi: 10.1007/s11747-019-00710-5.

[15] Sina Shafiezadeh, Gian Marco Duma, Giovanni Mento, Alberto Danieli, Lisa Antoniazzi, Fiorella Del Popolo Cristaldi, Paolo Bonanni, and Alberto Testolin. Methodological Issues in Evaluating Machine Learning Models for EEG Seizure Prediction: Good Cross-Validation Accuracy Does Not Guarantee Generalization to New Patients. *Applied Sciences*, 13(7):4262, March 2023. ISSN 2076-3417. doi: 10.3390/app13074262.

[16] Syed Muhammad Usman, Muhammad Usman, and Simon Fong. Epileptic Seizures Prediction Using Machine Learning Methods. *Computational and Mathematical Methods in Medicine*, 2017:1–10, 2017. ISSN 1748-670X, 1748-6718. doi: 10.1155/2017/9074759.

[17] Emina Alickovic, Jasmin Kevric, and Abdulhamit Subasi. Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. *Biomedical Signal Processing and Control*, 39:94–102, January 2018. ISSN 17468094. doi: 10.1016/j.bspc.2017.07.022.

[18] *K*ostas *M*. Tsiouris, Vasileios C. Pezoulas, Michalis Zervakis, Spiros Konitsiotis, Dimitrios D. Koutsouris, and Dimitrios I. Fotiadis. A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals. *Computers in Biology and Medicine*, 99:24–37, August 2018. ISSN 0010-4825. doi: 10.1016/j.compbiomed.2018.05.019.

[19] Levin Kuhlmann, Philippa Karoly, Dean R Freestone, Benjamin H Brinkmann, Andriy Temko, Alexandre Barachant, Feng Li, Gilberto Titericz, Brian W Lang, Daniel Lavery, Kelly Roman, Derek Broadhead, Scott Dobson, Gareth Jones, Qingnan Tang, Irina Ivanenko, Oleg Panichev, Timothée Proix, Michal Náhlík, Daniel B Grunberg, Chip Reuben, Gregory Worrell, Brian Litt, David T J Liley, David B Grayden, and Mark J Cook. Epilepsyecosystem.org: Crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain*, August 2018. ISSN 0006-8950, 1460-2156. doi: 10.1093/brain/awy210.

[20] Haidar Khan, Lara Marcuse, Madeline Fields, Kalina Swann, and Bulent Yener. Focal Onset Seizure Prediction Using Convolutional Networks. *IEEE Transactions on Biomedical Engineering*, 65(9):2109–2118, September 2018. ISSN 0018-9294, 1558-2531. doi: 10.1109/TBME.2017.2785401.

[21] Nhan Duy Truong, Anh Duy Nguyen, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Samuel Ippolito, and Omid Kavehei. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Networks*, 105:104–111, 2018. ISSN 0893-6080. doi: 10.1016/j.neunet.2018.04.018.

[22] Hisham Daoud and Magdy A. Bayoumi. Efficient Epileptic Seizure Prediction Based on Deep Learning. *IEEE Transactions on Biomedical Circuits and Systems*, 13(5):804–813, October 2019. ISSN 1932-4545, 1940-9990. doi: 10.1109/TBCAS.2019.2929053.

[23] Yankun Xu, Jie Yang, Shiqi Zhao, Hemmings Wu, and Mohamad Sawan. An End-to-End Deep Learning Approach for Epileptic Seizure Prediction. In *2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, pages 266–270, August 2020. doi: 10.1109/AICAS48895.2020.9073988.

[24] Ranjan Jana and Imon Mukherjee. Deep learning based efficient epileptic seizure prediction with EEG channel optimization. *Biomedical Signal Processing and Control*, 68:102767, July 2021. ISSN 17468094. doi: 10.1016/j.bspc.2021.102767.

[25] Ramy Hussein, Soojin Lee, Rabab Ward, and Martin J. McKeown. Semi-dilated convolutional neural networks for epileptic seizure prediction. *Neural Networks*, 139:

212–222, July 2021. ISSN 08936080. doi: 10.1016/j.neunet.2021.03.008.

[26] Ziyu Wang, Jie Yang, and Mohamad Sawan. A Novel Multi-scale Dilated 3D CNN for Epileptic Seizure Prediction. In *2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, pages 1–4, Washington DC, DC, USA, June 2021. IEEE. ISBN 978-1-66541-913-0. doi: 10.1109/AICAS51828.2021.9458571.

[27] Yikai Gao, Xun Chen, Aiping Liu, Deng Liang, Le Wu, Ruobing Qian, Hongtao Xie, and Yongdong Zhang. Pediatric Seizure Prediction in Scalp EEG Using a Multi-Scale Neural Network With Dilated Convolutions. *IEEE Journal of Translational Engineering in Health and Medicine*, 10:1–9, 2022. ISSN 2168-2372. doi: 10.1109/JTEHM.2022.3144037.

[28] Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces. *Journal of Neural Engineering*, 15(5):056013, October 2018. ISSN 1741-2560, 1741-2552. doi: 10.1088/1741-2552/aace8c.

[29] Kai Keng Ang, Zhang Yang Chin, Haihong Zhang, and Cuntai Guan. Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pages 2390–2397, Hong Kong, China, June 2008. IEEE. ISBN 978-1-4244-1820-6. doi: 10.1109/IJCNN.2008.4634130.

[30] Yuan Zhang, Yao Guo, Po Yang, Wei Chen, and Benny Lo. Epilepsy Seizure Prediction on EEG Using Common Spatial Pattern and Convolutional Neural Network. *IEEE Journal of Biomedical and Health Informatics*, 24(2):465–474, February 2020. ISSN 2168-2194, 2168-2208. doi: 10.1109/JBHI.2019.2933046.

[31] Imene Jemal, Neila Mezghani, Lina Abou-Abbas, and Amar Mitiche. An Interpretable Deep Learning Classifier for Epileptic Seizure Prediction Using EEG Data. *IEEE Access*, 10:60141–60150, 2022. ISSN 2169-3536. doi: 10.1109/ACCESS.2022.3176367.

[32] Ali Hossam Shoeb. *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*. PhD thesis, Massachusetts Institute of Technology, 2009.

[33] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), June 2000. ISSN 0009-7322, 1524-4539. doi: 10.1161/01.CIR.101.23.e215.

[34] Mohamad Shahbazi and Hamid Aghajan. A Generalizable Model for Seizure Prediction Based on Deep Learning Using Cnn-Lstm Architecture. In *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 469–473, Anaheim, CA, USA, November 2018. IEEE. ISBN 978-1-72811-295-4. doi: 10.1109/GlobalSIP.2018.8646505.

[35] Ahmet Remzi Ozcan and Sarp Erturk. Seizure Prediction in Scalp EEG Using 3D Convolutional Neural Networks With an Image-Based Approach. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(11):2284–2293, November 2019. ISSN 1534-4320, 1558-0210. doi: 10.1109/TNSRE.2019.2943707.

[36] Xiaoyan Wei, Lin Zhou, Zhen Zhang, Ziyi Chen, and Yi Zhou. Early prediction of epileptic seizures using a long-term recurrent convolutional network. *Journal of Neuroscience Methods*, 327:108395, November 2019. ISSN 01650270. doi: 10.1016/j.jneumeth.2019.108395.

[37] Jasmin Kevric and Abdulhamit Subasi. Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. *Biomedical Signal Processing and Control*, 31:398–406, January 2017. ISSN 17468094. doi: 10.1016/j.bspc.2016.09.007.

[38] Yves Meyer and David H. Salinger. *Wavelets and Operators*. Number 37 in Cambridge Studies in Advanced Mathematics. Cambridge university press, Cambridge, 1992. ISBN 978-0-521-42000-6.

[39] Serkan Kiranyaz, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J. Inman. 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151:107398, April 2021. ISSN 08883270. doi: 10.1016/j.ymssp.2020.107398.

[40] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature14539.