An Efficient Hybrid Model for Patient-Independent Seizure Prediction Using Deep Learning

Rowan Ihab Halawa *, Sherin M. Youssef and Mazen Nabil Elagamy *

Computer Engineering Department, College of Engineering and Technology, Arab Academy for Science and Technology (AAST), Alexandria 1029, Egypt; sherin@aast.edu
* Correspondence: rowanihab2210@gmail.com (R.I.H.); mazenelagamy@aast.edu (M.N.E.)

Abstract: Recently, many researchers have deployed different deep learning techniques to predict epileptic seizure, using electroencephalogram signals. However, most of this research requires very large amounts of memory and complicated feature extraction algorithms. In addition, they could not precisely examine EEG signal characteristics, which led to poor prediction performance. In this research, a non-patient-specific epileptic seizure prediction approach is proposed. The proposed model integrates Wavelet-based EEG signal processing with deep learning architectures for efficient prediction of pre-ictal and inter-ictal signals. The proposed system uses different models of one-dimensional convolutional neural networks to discriminate between inter-ictal signal and pre-ictal signals in order to enhance prediction performance. Experiments have been carried out on a benchmark dataset to validate the robustness of the proposed model. The experimental results showed that the proposed approach achieved 93.4% for 16 patients and 97.87% for 6 patients. Experiments showed that the proposed model can predict epileptic seizures effectively, which can have remarkable potential in clinical applications.

Keywords: epilepsy; detection; inter-ictal; ictal; pre-ictal; EEG signal; deep learning; 1-D CNN

1. Introduction

Epilepsy is a brain disease that causes sudden recurring seizures, which are provoked by unusual electrical discharges in some brain cells [1,2]. Such seizures can severely harm patients, as they might cause nervousness and injuries [3]. The World Health Organization (WHO) has stated that around 10% of the world population will have one seizure during their lifetime, which makes epilepsy one of the most common neurological disorders around the world [1–4]. Hence, finding an efficient methodology for epileptic seizure prediction is considered a very important issue [5]. Studies on seizure prediction using electroencephalogram (EEG) signals has shown that EEG holds valuable physiological and pathological information that can significantly enhance the performance of seizure predictive models if properly analyzed [1–7].

In addition, the use of deep learning (DL) and neural network (NNs) techniques in the development of computer models for diseases detection, the identification of diseases, and many other domains, have attracted the interest of numerous researchers [8–13]. For instance, the authors of [12] introduced a COVID-19 patient tracking system using Lagrange optimization and a distributed DL model to ensure that all data needed to track any suspected patient is received efficiently and reliably. To achieve a low square mean error, they deployed a one-dimensional convolution neural network (1D-CNN). Moreover, in real-world applications, such as that proposed in [13], a 1D-CNN deep learning model was used to calculate the optimum interference transmission power in communication between vehicles and everything. In addition, the authors of [14] showed the utility of DL for vehicle pollution detection using real-time images. Their outcome clearly demonstrated that a simple deep neural network (DNN) architecture is capable of producing significant
enhancements in results. Likewise, the authors of [15] proposed a DL model based on a convolutional neural network (CNN) and spectrogram images from EEG signals in major depression detection. Meanwhile, in [16], to classify left-hand and right-hand movements from motor imagery–EEG signals, a classification model (MIDNN) based on the DNN model was developed. Using the Brain Computer Interface (BCI) and Emotiv datasets, their proposed model improved classification accuracy by 13.7 percent and 26.9 percent, respectively. The promising results achieved by computer models deploying DL and NNs in different domains have motivated other researchers to examine the effectiveness of using DL and NNs in analyzing epilepsy EEG signals.

Consequently, extensive studies using different DL and NNs techniques applied to EEG signals have been undertaken in order to support epileptic seizure detection and prediction. In 2020, the authors of [17] introduced a spectrogram and 1D-CNN-based approach to detect epileptic seizures using EEG signals. Their design deployed a spectrogram for feature extraction (FE), which was then used as an input for the 1D-CNN for epileptic seizures detection. Using the Children's Hospital Boston and the Massachusetts Institute of Technology (CHB-MIT) datasets as a benchmark, their approach only achieved 77.57% accuracy in seizure detection [17]. Hence, the authors of [18] used sample entropy as a feature to feed to the Bi-LSTM model for seizure prediction using CHB-MIT dataset achieving an accuracy of 80.09% working on 13 patients. Later, in October 2020, the authors of [19] presented a binary single-dimensional convolutional neural network (BSDCNN) for seizure prediction that was tested using two benchmarks, the American Epilepsy Society Seizure Prediction Challenge (AES) dataset and the CHB-MIT. Their design diminished the parameter size by 7.2 times and the computation complexity by 25.5 times while achieving a precision loss of 2.74% [19]. Furthermore, the authors of [20] proposed a one-dimensional capsule networks model for pre-ictal/inter-ictal perception, which reached a 97.74% classification accuracy using the F3-C3 channel pair. The experimental results showed that the 30 min prior to a seizure holds very valuable information that can support seizure prediction [20]. Recently, for implementing a patient-specific seizure prediction system, the authors of [21] designed an end-to-end framework that uses dilated convolutions with a multi-scale convolutional neural network. Their model reached an average proportion of time-in-warning of 6.3%, an average sensitivity of 93.3%, and an average false prediction rate of 0.007 per hour when tested on 16 patients from the CHB-MIT dataset [21]. Moreover, the authors of [22] designed a custom spiking neural network (SNN) to detect events of interest (EoI) in the 80–250 Hz ripple band and reject artifacts in the 500–900 Hz band. Meanwhile, the authors of [23] proposed a patient-specific prediction model. They converted raw signals into images and reached an accuracy of 94.3% when seizure prediction was done for five patients of the CHB-MIT dataset.

However, the conversion of EEG signals to images could cause the loss of important information that CNNs could extract from raw signals. For the purpose of noise removal from EEG signals, discrete wavelet transform (DWT) has been widely used. The Daubechies sequence has good scalability and flexibility for the evaluation of boundary issues, according to [24]. In addition, Daubechies-4 (db4) and Daubechies-6 (db6) wavelets were used to eliminate noise from EEG signals in [9] and [10], respectively. In addition, in [25], it was proved that “Symlet9” was the most suitable wavelet transform for EEG signals to achieve wavelet de-noising through decomposition and reconstruction.

Despite the fact that many researchers have addressed the issue of predicting epileptic seizures using DL learning techniques, the majority of these studies were patient-specific and relied on the historical data of a few patients to train their models. In addition, most previous research did not properly tackle the issue of the noisy channels of EEG signals. Being that they are scalp signals, EEG signals are challenging due to the presence of noise, which requires a pre-processing step before the prediction step. Consequently, this research focuses on determining if analyzing the historical data of all patients in the CHB-MIT dataset will support designing an efficient patient independent epileptic seizures prediction model. Additionally, a noise removal module is introduced using DWT techniques.
The aim of this research is to design and implement an efficient hybrid model for the prediction of epileptic seizures using DL techniques. In addition, it aims to study the effect of using DWT techniques to introduce a noise removal module on the performance of the presented model.

The contributions of this paper can be demonstrated as follows. A generalized non-patient-specific epileptic seizure prediction model is developed. The proposed model integrates DWT for EEG signal processing with DL techniques for efficient classification of EEG signals. Furthermore, it introduces a robust noise removal technique that is adaptable to different EEG signal channel selections.

The rest of this paper is organized as follows. The next section describes the materials and methods used, including the dataset, explains the different modules used in the proposed approach and shows the 1-D CNN models’ architectures. In Section 3 the experimental results using the two proposed DL models are presented. The analysis is then discussed in Section 4. Finally, Section 5 presents the conclusions.

2. Materials and Methods

The dataset used in this research was assembled from the CHB-MIT [26], which is a benchmark dataset and is freely accessible for research.

Each case (chb1, chb2, etc.) contained between 9 and 42 continuous `.edf` files from a single patient. To protect the patients’ privacy, all protected health information (PHI) in the original files was replaced with surrogate information in the provided files. Surrogate dates replaced the dates in the original files, but the time relationships between the individual files belonging to each case were preserved. The files, in most cases, contained exactly one hour of digitized EEG signals, though those in case chb10 were two hours long, and those in cases chb04, chb06, chb07, chb09, and chb23 were four hours long; files in which seizures were recorded were occasionally shorter. For these recordings, the International 10–20 system of EEG electrode positions and nomenclature was used.

Files labeled ‘RECORDS’ contained a list of all 664 files in the collection, while files labeled ‘RECORDS-WITH-SEIZURES’ contained a list of the 129 files that contained one or more seizures. These records contained 198 seizures in total, with the beginning and the end of each seizure annotated.

This dataset included continual EEG signals recordings for 24 patients, composed of 5 males and 17 females patients. The records for patient 1 and 21 were gathered from the same female with 1.5-year difference. While records for patient 24 were added to the dataset later in December 2010. Female patients’ ages were from 1.5 to 19 years, while male patients’ ages varied from 3 to 22 years. The data were reported and collected by positioning 23 electrodes on the scalps of patients, which meant that all EEG signals were scalp EEG (sEEG). This dataset has a frequency of 256 HZ per second with a 16-bit resolution.

Figure 1 shows a sample of the dataset used. The signal depicted represents patient one’s ictal phase, which is the phase in which he had seizures that typically last a few seconds. The figure shows that, for each patient, there were several parallel channels, which could differ from one patient to the next. Every patient’s record file contained information about the exact location and duration of the ictal state. The pre-ictal state is the state that occurs immediately before the onset of the ictal phase, whereas the inter-ictal state was the state in which the patient was normal. To achieve standardization, the following 18 channels were chosen from the 23 electrodes to be the 18 common channels between all 24 patients: ['FP2-F4', 'F3-C3', 'P7-O1', 'P8-O2', 'T8-P8', 'FP1-F7', 'F8-T8', 'T7-P7', 'F4-C4', 'FP2-F8', 'CZ-PZ', 'C3-P3', 'F7-T7', 'C4-P4', 'FP1-F3', 'FZ-CZ', 'P3-O1', 'P4-O2'].

[Source: https://www.mdpi.com/2076-3417/12/14/5516]
Seizure prediction methods are divided into patient-specific and patient-independent. EEG signal patterns are unique for each epilepsy patient. Moreover, seizure prediction is a challenging task because of the presence of noise during EEG recording due to eye blinking, eyelids movement, swallowing, and breathing. In addition, EEG signal patterns vary from one individual to another.

The proposed system for epileptic prediction system, as shown in Figure 2, is composed of 3 main modules: EEG signal pre-processing, EEG signal segmentation, and the 1-D CNN based model used to predict whether the signal is pre-ictal or inter-ictal.

As shown in Figure 2, the signal channels are first pre-processed and noise is removed from every channel using DWT techniques. After channel cleaning, the pre-ictal and inter-ictal segments are extracted from the signals. Each channel is then passed to the 1-D CNN model for the prediction process.

2.1. EEG Signal Pre-Processing

The pre-processing step’s goal here is to remove noise from the EEG signal in order to get more accurate features and results. Many studies have discussed different noise removal techniques, and one of the most effective methods is discrete wavelet de-noising schema. Wavelet de-noising is effective in de-noising physiological signals because it has the tendency to preserve signal properties while reducing noise; this is preferred and frequently used over signal frequency domain filtering. The reason for this is that threshold strategies are available that allow reconstruction based on chosen coefficients. It has also frequently used over signal frequency domain filtering. The reason for this is that threshold strategies are available that allow reconstruction based on chosen coefficients. It has also

Figure 1. Sample of CHB-MIT dataset represented by parallel readings from electrodes placed on a patient’s scalp.

![Figure 1](image1.png)

Figure 2. Proposed model architecture, consisting of a pre-processing module followed by an EEG signal segmentation phase and a deep learning phase.

![Figure 2](image2.png)
strategies are available that allow reconstruction based on chosen coefficients. It has also been proven in recent years that the following sets of wavelets are effective for cleaning signals: ['Daub4', 'Daub6', 'Daub8', 'Daub12', 'DMeyer', 'Symlet9'].

In this paper, the latter ones were applied to choose the best de-noising technique for every EEG channel. The dataset signals were further subdivided into individual channels. Each of these channels was then fed into the noise removal module as an input. This module included six major wavelets, which were as follows: ['Daub4', 'Daub6', 'Daub8', 'Daub12', 'DMeyer', 'Symlet9']. The 'denoise_wavelet' method from the Python library 'skimage' was used for each of the wavelets, with wavelet levels = 4.

The wavelet domain is a sparse representation of the signal, similar to the frequency domain of the Fourier transform. Sparse representations have primarily values of zero or near-zero and truly random noise is represented by many small values in the wavelet domain. Setting all values below a certain threshold helps to reduce signal noise.

A de-noising procedure is usually divided into three steps. First, a wavelet and a N level are selected. The wavelet decomposition of signal s at level N is then computed. In the case of this study, N was set to 4, and we had a list of several wavelets, as previously mentioned. Figure 3a depicts the decomposition of a signal into different levels using low pass and high pass filters. Second, a threshold was chosen for each level, from 1 to N, and soft thresholding was applied to the detail coefficients. Finally, wavelet reconstruction was computed using the original approximation coefficients of level N and the modified detail coefficients of levels 1–N.

For each channel, the best wavelet was chosen according to the least root mean square error (RMSE) and the highest signal-to-noise ratio (SNR). SNR is defined as the ratio between the power of the signal and the power of the noise; this ratio must be maximized to reach a better and cleaner signal. While RMSE is the standard deviation of the residual (difference between the original signal and cleaned signal), it is used frequently in forecasting and regression analyses to verify experimental results. RMSE and SNR have been widely used as criteria to choose the most effective wavelet for a cleaned signal. Table 1 shows examples of different SNR and RMSE values for a channel of an EEG signal from patient one. Based on the experimental results, channel ‘db12’ had the highest SNR and the lowest RMSE, so it was chosen as the best wavelet for noise removal in this channel.
Table 1. Performance comparison of different wavelet families, applied to EEG channels, in terms of signal-to-noise ratio (SNR), and root mean square error (RMSE).

| Wavelet   | SNR   | RMSE  |
|-----------|-------|-------|
| ‘Daub4’   | 55.12 | 0.0036|
| ‘Daub6’   | 55.46 | 0.0035|
| ‘Daub8’   | 53.43 | 0.0040|
| ‘Daub12’  | 56.189| 0.0034|
| ‘DMeyer’  | 55.578| 0.0035|
| ‘Symlet9’ | 53.80 | 0.0038|

In order to prove the robustness of the system and the effectiveness of the noise removal module, artificial noise, such as Gaussian noise, was added to the signal at different percentages. In Figure 4, the average accuracy of the four classifiers (SVM, random forest, KNN, and bagging classifier) is presented against different percentages of white Gaussian noise. This showed that the system achieved a high accuracy, even with added noise. We could conclude from this experiment that the system was almost robust up to 60% noise as the accuracy was still above 95%.

Figure 4. White Gaussian noise percentage vs. average accuracy over SVM, RF, KNN and bagging classifier.

2.2. EEG Signal Segmentation Phase

After noise removal in all channels for each patient, the following segments were extracted:

1. Pre-ictal phase segment (with window = k), which represents the window of signal that happens just before the ictal phase and which could be used as an alarm to take precautions by caregivers.
2. Inter-ictal phase segment (with window = k), which represents the normal signal where the patient is in normal state.

The ‘k’ value was tested with different values to study the effect of the segmentation of the signal. First k was equal to 8 minutes for both the pre-ictal and inter-ictal phases. Furthermore, the window extracted was segmented to equally lengthen the windows of L length, where L was chosen to be 10 s and 30 s. Both overlapping and non-overlapping windows were considered in the segmentation phase.

Figure 5 shows the channel signal after extracting the inter-ictal phase, where k = 8 min.
Figure 5. The orange part of the signal represents k minutes of the inter-ictal phase.

2.3. EEG Signal Prediction Using Deep Learning Models

In this paper, two DL models, both based on the 1D CNN, are presented. The common EEG channels between all subjects in the dataset were each passed using the proposed models and with a majority voting scheme, the results were reached as a prediction of the given signal. Figure 6 shows the proposed diagram of the deep learning model. Each channel is passed as an input, independently of the other channels, to the CNN model.

Figure 6. Proposed deep learning method.

2.4. Layers of CNN Model Architecture

- **Convolution Layer:** In this layer, the input dataset is convolved with a kernel or weight. A feature map is considered the output of this layer. The stride parameter determines how much the kernel convolves with the input data. By learning from the diverse input signals, this operation is considered as a feature extractor.

  \[
  \text{Conv1D} \left( \text{filters, kernel-size, strides, activation} = \text{none} \right)
  \]

  - Filters: Number of sliding windows;
  - Kernel-size: The size of the sliding window;
The next layer is BatchNormalization, followed by a dropout layer, and finally a Sigmoid layer. Hence, three blocks were adopted in the proposed model.

Experimental results showed that the performance increased when block two and then block three were added. Adding more than three blocks began to reduce the model performance. The effect of adding blocks two and three on the performance were examined. The experimental results showed that the performance increased when block two and then block three were added. Adding more than three blocks began to reduce the model performance.

The first proposed model architecture is presented in Figure 7. The model is structured as three main CNN blocks, each of which includes four layers, 1D CNN, Max Pool 1D, BatchNormalization, and spatial dropout. After the end of the third CNN block, the output is passed to a flatten layer and then three dense layers with the activation function ‘relu’. The presence of more than one dense layer aids in improving the model’s performance. The next layer is BatchNormalization, followed by a dropout layer, and finally a Sigmoid layer for classification of the signal as either inter-ictal or pre-ictal. The main goal of the layers after the third block is to avoid overfitting of the data.

2.5. First Proposed Model Architecture

The first proposed model architecture is presented in Figure 7. The model is structured as three main CNN blocks, each of which includes four layers, 1D CNN, Max Pool 1D, BatchNormalization, and spatial dropout. After the end of the third CNN block, the output is passed to a flatten layer and then three dense layers with the activation function ‘relu’. Overfitting is avoided by setting input units to zero at random with a rate of frequency at each step during the training period. Inputs that are not set to zero are scaled up by 1/(1-rate) so that the total sum remains the same for all inputs.

Pooling layer: This is a down-sampling layer that uses the pooling process to lower the input data’s dimension (reduce the amount of computation performed in the network) and increase the robustness of feature extraction while keeping all the important information. Average, max, and sum are all considered as pooling layer operations. The max-pooling operation is used more often in modern deep learning architecture.

- Pool-size: take the maximum value in a defined spatial window.
- Batch normalization: A method for training deep neural networks that normalizes each contribution to a layer for every small batch. This slows down the learning process and reduces the number of training epochs required to train and build deep neural networks.
- Dropout: Overfitting is avoided by setting input units to zero at random with a rate of frequency at each step during the training period. Inputs that are not set to zero are scaled up by 1/(1-rate) so that the total sum remains the same for all inputs.

Pooling layer:
- MaxPooling1D (pool-size = 2)
- Pool-size: take the maximum value in a defined spatial window.
- Batch normalization: A method for training deep neural networks that normalizes each contribution to a layer for every small batch. This slows down the learning process and reduces the number of training epochs required to train and build deep neural networks.
- Dropout: Overfitting is avoided by setting input units to zero at random with a rate of frequency at each step during the training period. Inputs that are not set to zero are scaled up by 1/(1-rate) so that the total sum remains the same for all inputs.

2.5. First Proposed Model Architecture

The first proposed model architecture is presented in Figure 7. The model is structured as three main CNN blocks, each of which includes four layers, 1D CNN, Max Pool 1D, BatchNormalization, and spatial dropout. After the end of the third CNN block, the output is passed to a flatten layer and then three dense layers with the activation function ‘relu’. The presence of more than one dense layer aids in improving the model’s performance. The next layer is BatchNormalization, followed by a dropout layer, and finally a Sigmoid layer for classification of the signal as either inter-ictal or pre-ictal. The main goal of the layers after the third block is to avoid overfitting of the data.

2.5. First Proposed Model Architecture

The first proposed model architecture is presented in Figure 7. The model is structured as three main CNN blocks, each of which includes four layers, 1D CNN, Max Pool 1D, BatchNormalization, and spatial dropout. After the end of the third CNN block, the output is passed to a flatten layer and then three dense layers with the activation function ‘relu’. The presence of more than one dense layer aids in improving the model’s performance. The next layer is BatchNormalization, followed by a dropout layer, and finally a Sigmoid layer for classification of the signal as either inter-ictal or pre-ictal. The main goal of the layers after the third block is to avoid overfitting of the data.

Figure 7. First proposed deep learning model architecture.

The model was initially constructed with just one block, containing four layers, 1-D CNN, maxPool 1D, BatchNormalization, and spatial dropout. During the experiments, the effect of adding blocks two and three on the performance were examined. The experimental
results showed that the performance increased when block two and then block three were added. Adding more than three blocks began to reduce the model performance. Hence, three blocks were adopted in the proposed model.

2.6. Second Proposed Model Architecture

The second proposed model architecture is presented in Figure 8. The model was structured as four main CNN blocks, each of which included four layers: two layers of 1D CNN followed by Max Pool 1D, and spatial dropout. After the end of the fourth CNN block, the output is passed to a dropout layer and finally a Sigmoid layer for classification of the signal as inter-ictal or pre-ictal. This model is considered different from the first model and the goal was to compare and study the different effects of varying the deep learning model on the results. The difference was that each block has one more layer of 1-D CNN and the dropout occurs only once after the four blocks. A fourth block was also added to this model to make it complex enough for the EEG signal prediction process.

Data augmentation has been integrated in the model to guarantee a big enough dataset for the deep learning models which always have a need for larger datasets in order to achieve two main goals: the first is to avoid overfitting and the second is to achieve high performance results in validation and testing. In this paper, data augmentation was applied using the noise addition technique.

3. Results

Both models were trained using various batch sizes. The dataset was divided into 70 percent for training and 30 percent for testing. To effectively assess the predictive power of the model, the training dataset was further split, and 20 percent of the training dataset was utilized for inclusive validation. Different types of optimizers were tested, such as ‘SGD’ and ‘Adam’. Based on experimental analysis, the ‘Adam’ optimizer outperformed other optimizers by a wide margin by providing an optimized gradient descent. Hence, ‘Adam’ was chosen as a suitable optimizer for the proposed model.

In the following tables (Tables 2–7), a sample of the results is presented for model 1 and model 2, with different k window values, where k = 8 min, 30 s, and 10 s, in terms of accuracy, area under the curve, precision, and recall. It is clear from the results presented in these tables that model 1 outperforms model 2. Furthermore, as shown in the tables, the model performs best for k = 10 s, followed by k = 30 s, and finally for k = 8 min.
Table 2. Accuracy, area under curve (AUC), precision and recall results for model 1, 8 min window.

| Channel | AUC (%) | Accuracy (%) | Precision (%) | Recall (%) |
|---------|---------|--------------|---------------|------------|
| Channel 1 | 85% | 77.6% | 74.7% | 80% |
| Channel 2 | 88% | 78.1% | 77.7% | 72.4% |
| Channel 3 | 84% | 73.8% | 68.97% | 80% |
| Channel 4 | 85.5% | 78.5% | 72% | 88% |
| Channel 5 | 90% | 75.9% | 58% | 82% |
| Channel 6 | 84.4% | 75.6% | 79% | 69.9% |

Table 3. Accuracy, area under curve (AUC), precision and recall results for model 1, 30 s epochs.

| Channel | AUC (%) | Accuracy (%) | Precision (%) | Recall (%) |
|---------|---------|--------------|---------------|------------|
| Channel 1 | 85.6% | 84.78% | 83.47% | 81% |
| Channel 2 | 89.76% | 82.97% | 82.57% | 83.25% |
| Channel 3 | 86% | 84.6% | 78% | 83.6% |
| Channel 4 | 87.6% | 85.5% | 83.7% | 89.7% |
| Channel 5 | 92.9% | 83.13% | 79.79% | 87% |
| Channel 6 | 89.4% | 85.95% | 81% | 71% |

Table 4. Accuracy, area under curve (AUC), precision and recall results for model 1, 10 s epochs.

| Channel | AUC (%) | Accuracy (%) | Precision (%) | Recall (%) |
|---------|---------|--------------|---------------|------------|
| Channel 1 | 92.76% | 85.97% | 86% | 84.3% |
| Channel 2 | 89% | 86.6% | 78% | 83.6% |
| Channel 3 | 90.6% | 89% | 83.7% | 89.7% |
| Channel 4 | 92.9% | 86.13% | 79.79% | 87% |
| Channel 5 | 89.9% | 87.16% | 84% | 88.5% |
| Channel 6 | 88.3% | 84.16% | 83.3% | 73.3% |

Table 5. Accuracy, area under curve (AUC), precision and recall results for model 2, 8 min window.

| Channel | AUC (%) | Accuracy (%) | Precision (%) | Recall (%) |
|---------|---------|--------------|---------------|------------|
| Channel 1 | 71% | 66% | 67.5% | 64% |
| Channel 2 | 73.5% | 76.6% | 80% | 68% |
| Channel 3 | 80% | 66.36% | 60.29% | 82% |
| Channel 4 | 76.5% | 68% | 61.24% | 79% |
| Channel 5 | 78.76% | 72.56% | 70% | 79% |
| Channel 6 | 74.9% | 71.32% | 76.8% | 80% |
Table 6. Accuracy, area under curve (AUC), precision and recall results for model 2, 30 s epochs.

| Channel | AUC (%) | Accuracy (%) | Precision (%) | Recall (%) |
|---------|---------|--------------|---------------|------------|
| Channel 1 | 78.38% | 72.2% | 72.75% | 66.9% |
| Channel 2 | 84.93% | 80% | 79.59% | 78% |
| Channel 3 | 77.8% | 76% | 73% | 74% |
| Channel 4 | 81.37% | 70.29% | 68.88% | 88% |
| Channel 5 | 84.47% | 78.5% | 72.13% | 88% |
| Channel 6 | 84.45% | 75.67% | 79% | 81.15% |

Table 7. Accuracy, area under curve (AUC), precision and recall results for model 2, 10 s epochs.

| Channel | AUC (%) | Accuracy (%) | Precision (%) | Recall (%) |
|---------|---------|--------------|---------------|------------|
| Channel 1 | 85.76% | 78.43% | 78% | 79.8% |
| Channel 2 | 83% | 82.27% | 76% | 74.68% |
| Channel 3 | 85.8% | 80% | 79% | 81% |
| Channel 4 | 87% | 79.79% | 72% | 90% |
| Channel 5 | 84.47% | 84.3% | 72.13% | 88% |
| Channel 6 | 86.5% | 78.54% | 82% | 86.3% |

In Figure 9, examples of the validation vs. training accuracy are presented. A comparison between the results for testing, validation, and training is described in Tables 8 and 9.

Table 8. Accuracy results for model 1 (training vs. validation vs. testing).

|             | Accuracy (%) |
|-------------|--------------|
| 8 min       | 30 s         | 10 s         |
| Train       | 96%          | 97%          | 98%          |
| Validation  | 85%          | 88%          | 90%          |
| Testing     | 76%          | 84.45%       | 87%          |

Table 9. Accuracy results for model 2 (training vs. validation vs. testing).

|             | Accuracy (%) |
|-------------|--------------|
| 8 min       | 30 s         | 10 s         |
| Train       | 94.4%        | 96%          | 96.89%       |
| Validation  | 84%          | 85.79%       | 87%          |
| Testing     | 70%          | 75.35%       | 80%          |
Figure 9. (a) Training vs. validation accuracy for channel 1, for model 2, where window = 10 s. (b) Training vs. validation accuracy for channel 1, for model 1, where window = 10 s. (c) Training vs. validation accuracy for channel 6, for model 2, where window = 30 s. (d) Training vs. validation accuracy for channel 5, for model 2, where window = 30 s.

4. Discussion

It is notable from the previous section that there were promising results for both models 1 and 2 with the DL approach. The proposed model can help medical systems to predict seizures in patients with a previous medical history, as well as patients lacking a medical history.

The work presented in this paper shows the great effect on results when segmenting data into smaller epochs in order to enrich the data fed into the DL model. In Figure 10, the effect of segmentation on the proposed model 1, using a window of $L$ length, where $L = 10$, 30 s, and 8 min are shown.

A comparative study was undertaken using prior related work, using the same benchmark dataset (CHB-MIT) and the results are presented in Table 10. As shown in the table, the proposed work outperforms state-of-the-art methods in seizure prediction by testing a larger number of samples and by being a patient-independent model. In [19], a binary single-dimensional CNN (BSDCNN) for epileptic seizure prediction was proposed. Their model uses 1D convolutional kernels in the prediction process. Furthermore, the authors of [20] implemented a one-dimensional capsule network for epileptic seizure prediction that achieved an accuracy of 97.74% using only two EEG signal channels. In addition, in [21], a patient-specific prediction model that uses a temporal-spatial multi-scale CNN was presented. Working with 16 patients, they achieved an accuracy of 93.3%. Finally, a patient-specific epileptic seizure prediction model was presented in [23]. They converted raw signals into images. Their model reached an accuracy of 94.3% when seizure prediction was done for 5 patients. However, the conversion of EEG signals into images could cause the loss of important information that CNN can extract from raw signals. Results showed that the proposed model performed better in different cases, such as when using 6 patients
and 10 patients for prediction using DL. The improvements in rates were 0.87% compared to the values in [19], 0.13% compared to values in [20], 0.1% compared to values in [21], and 3.54% compared to the values in [23] respectively. The work presented in this paper provides a generic model that can use patients’ historical data to predict seizures for other new patients with a high accuracy and precision.

![Figure 10. Comparison of average accuracy for different k length windows.](image)

Table 10. Comparative analysis study with previous state-of-the-art work in the same domain.

| Work               | Approach/Dataset          | Method              | Accuracy (%) |
|--------------------|---------------------------|---------------------|--------------|
| Zhao et al. [19]   | Patient specific/CHB-MIT. 6 patients. | 1D CNN based model | 97%          |
| Toraman et al. [20]| Patient specific/CHB-MIT. 2 channels only. | 1D CNN based model | 97.74%       |
| Gao et al. [21]    | Patient specific/CHB-MIT. 16 patients only. | 1D CNN based model | 93.3%        |
| Jana et al. [23]   | Patient specific/CHB-MIT. 5 patients. | CNN based model     | 94.33%       |
| Proposed work      | Non-patient specific/CHB-MIT. 6 patients. | 1D CNN based model | 97.87%       |
| Proposed work      | Non-patient specific/CHB-MIT. 16 patients. | 1D CNN based model | 93.4%        |

5. Conclusions

A generalized epileptic seizure hybrid prediction model using DL techniques was introduced to minimize risks that can affect the patients during a seizure. The proposed model initially performed pre-processing on EEG signals, using the DWT technique, in order to acquire a noise-free signal, followed by deploying segmentation of the channel signal to different k-window length segments in order to enhance the performance of epileptic seizure prediction. The proposed model achieved an accuracy of 97.87% for 6 patients and 93.4% for 16 patients, respectively. Achieving improvement rates of 0.87% compared to the work in [19], 0.13% compared to those of [20], 0.1% compared to values in [21], and 3.54% compared to those of [23], respectively. These results show that the proposed work enhances the performance of epileptic seizure prediction through examining EEG signals with a high accuracy compared to some of state-of-the-art models in the literature. The findings in this paper could help to save patients’ lives, learn more about the nature of pre-ictal vs. inter-ictal signals, and support patients to take important precautions to reduce risks during seizures.

The limitation of the proposed work is that only two categories of seizure, pre-ictal and inter-ictal, are analyzed. Further work can be extended to deal with other categories of
seizures. In addition, further research can examine the integration of the proposed model with cloud services and to extend the model for application to IoT platforms.

**Author Contributions:** Conceptualization, S.M.Y. and M.N.E.; methodology, S.M.Y. and M.N.E.; software, R.I.H.; validation, R.I.H., S.M.Y. and M.N.E.; formal analysis, R.I.H., S.M.Y. and M.N.E.; investigation, R.I.H., S.M.Y. and M.N.E.; resources, R.I.H. and M.N.E.; data curation, R.I.H.; writing—original draft preparation, R.I.H. and M.N.E.; writing—review and editing, R.I.H., S.M.Y. and M.N.E.; visualization, R.I.H., S.M.Y. and M.N.E.; supervision, S.M.Y. and M.N.E.; project administration, S.M.Y. and M.N.E.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available in a publicly accessible repository. Publicly available datasets were analyzed in this study. This data can be found here: [https://physionet.org/content/chbmit/1.0.0/](https://physionet.org/content/chbmit/1.0.0/) (accessed on 30 April 2022) and [https://www.upf.edu/web/ntsa/downloads](https://www.upf.edu/web/ntsa/downloads), (accessed on 30 April 2022). The dataset files’ license can be found at [https://physionet.org/content/chbmit/view-license/1.0.0/](https://physionet.org/content/chbmit/view-license/1.0.0/) (accessed on 30 April 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Jiang, Y.; Chen, W.; Li, M. Symplectic geometry decomposition-based features for automatic epileptic seizure detection. *Comput. Biol. Med.* 2020, 116, 103549. [CrossRef] [PubMed]

2. Yuan, Y.; Xun, G.; Jia, K.; Zhang, A. A Multi-View Deep Learning Framework for EEG Seizure Detection. *IEEE J. Biomed. Health Inform.* 2018, 23, 83–94. [CrossRef] [PubMed]

3. Racine, R.J. Modification of seizure activity by electrical stimulation: II. Motor seizure. *Electroencephalogr. Clin. Neurophysiol.* 1972, 32, 281–294. [CrossRef]

4. World Health Organization, Epilepsy. Available online: [https://www.who.int/en/news-room/fact-sheets/detail/epilepsy](https://www.who.int/en/news-room/fact-sheets/detail/epilepsy) (accessed on 15 April 2022).

5. Mirovska, P.; Madhavan, D.; LeCun, Y.; Kuzniecky, R. Classification of patterns of EEG synchronization for seizure prediction. *Clin. Neurophysiol.* 2009, 120, 1927–1940. [CrossRef] [PubMed]

6. Vidyaratne, L.S.; Iftekharuddin, K.M. Real-Time Epileptic Seizure Detection Using EEG. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2017, 25, 2146–2156. [CrossRef] [PubMed]

7. Sharma, M.; Shah, S.; Achuth, P.V. A novel approach for epilepsy detection using time–frequency localized bi-orthogonal wavelet filter. *J. Mech. Med. Biol.* 2019, 19, 1940007. [CrossRef]

8. Ashokkumar, S.R.; Anupallavi, S.; Premkumar, M.; Jeevanantham, V. Implementation of deep neural networks for classifying electroencephalogram signal using fractional S-transform for epileptic seizure detection. *Int. J. Imaging Syst. Technol.* 2021, 31, 895–908. [CrossRef]

9. Gao, Y.; Gao, B.; Chen, Q.; Liu, J.; Zhang, Y. Deep Convolutional Neural Network-Based Epileptic Electroencephalogram (EEG) Signal Classification. *Front. Neurol.* 2020, 11, 375. [CrossRef]

10. Salamatian, A.; Khadem, A. Automatic Sleep Stage Classification Using 1D Convolutional Neural Network. *Front. Biomed. Technol.* 2020, 7, 142–150. [CrossRef]

11. Giudice, M.L.; Varone, G.; Ieracitano, C.; Mamnone, N.; Bruna, A.R.; Tomaselli, V.; Morabito, F.C. 1D Convolutional Neural Network approach to classify voluntary eye blinks in EEG signals for BCI applications. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; IEEE: New York, NY, USA, 2020; pp. 1–7.

12. Osman, R.A.; Saleh, S.N.; Saleh, Y.N.M.; Elagamy, M.N. A Reliable and Efficient Tracking System Based on Deep Learning for Monitoring the Spread of COVID-19 in Closed Areas. *Int. J. Environ. Res. Public Health* 2021, 18, 12941. [CrossRef]

13. Osman, R.A.; Saleh, S.N.; Saleh, Y.N.M.; Elagamy, M.N. Enhancing the Reliability of Communication between Vehicle and Everything (V2X) Based on Efficient Road Traffic Information. *Appl. Sci.* 2021, 11, 11382. [CrossRef]

14. Cui, J.; Lan, Z.; Sourina, O.; Müller-Wittig, W. EEG-Based Cross-Subject Driver Drowsiness Recognition With an Interpretable Convolutional Neural Network. *IEEE Trans. Neural Netw. Learn. Syst.* 2022, 1–13. [CrossRef] [PubMed]

15. Loh, H.W.; Ooi, C.P.; Aydemir, E.; Tuner, T.; Dogan, S.; Acharya, U.R. Decision support system for major depression detection using spectrogram and convolution neural network with EEG signals. *Expert Syst.* 2022, 39, e12773. [CrossRef]

16. Jana, G.C.; Sharma, R.; Agrawal, A. A 1D-CNN-Spectrogram Based Approach for Seizure Detection from EEG Signal. *Procedia Comput. Sci.* 2020, 167, 403–412. [CrossRef]
18. Zhang, Q.; Ding, J.; Kong, W.; Liu, Y.; Wang, Q.; Jiang, T. Epilepsy prediction through optimized multidimensional sample entropy and Bi-LSTM. *Biomed. Signal Process. Control.* 2020, 64, 102293. [CrossRef]

19. Zhao, S.; Yang, J.; Xu, Y.; Sawan, M. Binary single-dimensional convolutional neural network for seizure prediction. In Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Sevilla, Spain, 10–21 October 2020; IEEE: New York, NY, USA, 2020; pp. 1–5. [CrossRef]

20. Toraman, S. Automatic recognition of preictal and interictal EEG signals using 1D-capssule networks. *Comput. Electr. Eng.* 2021, 91, 107033. [CrossRef]

21. Gao, Y.; Chen, X.; Liu, A.; Liang, D.; Wu, L.; Qian, R.; Xie, H.; Zhang, Y. Pediatric Seizure Prediction in Scalp EEG Using a Multi-Scale Neural Network With Dilated Convolutions. *IEEE J. Transl. Eng. Health Med.* 2022, 10, 1–9. [CrossRef]

22. Burelo, K.; Ramantani, G.; Indiveri, G.; Sarnthein, J. A neuromorphic spiking neural network detects epileptic high frequency oscillations in the scalp EEG. *Sci. Rep.* 2022, 12, 1798. [CrossRef]

23. Jana, R.; Bhattacharyya, S.; Das, S. Patient-specific seizure prediction using the convolutional neural networks. In *Intelligence Enabled Research*; Springer: Singapore, 2020; pp. 51–60.

24. Aliyu, I.; Lim, C.G. Selection of optimal wavelet features for epileptic EEG signal classification with LSTM. *Neural Comput. Appl.* 2021, 1–21. [CrossRef]

25. Al-Qazzaz, N.K.; Bin Mohd Ali, S.H.; Ahmad, S.A.; Islam, M.S.; Escudero, J. Selection of mother wavelet functions for multichannel EEG signal analysis during a working memory task. *Sensors* 2015, 15, 29015–29035. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4701319/ (accessed on 20 May 2022). [CrossRef] [PubMed]

26. EEG Time Series Data (Children’s Hospital Boston and the Massachusetts Institute of Technology). Available online: https://physionet.org/physiobank/database/chbmit/ (accessed on 7 May 2022).