Article

Epileptic Seizure Detection from Decomposed EEG Signal through 1D and 2D Feature Representation and Convolutional Neural Network

Shupta Das [1], Suraiya Akter Mumu [1], M. A. H. Akhand [1,*], Abdus Salam [2] and Md Abdus Samad Kamal [3,*]

1 Department of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna 9203, Bangladesh; shuptadas25203@gmail.com (S.D.); suraiyaaktermumu26@gmail.com (S.A.M.)
2 Shaheed Sheikh Abu Naser Specialized Hospital, Khulna 9000, Bangladesh; a.salamk51@gmail.com
3 Graduate School of Science and Technology, Gunma University, Kiryu 376-8515, Japan
* Correspondence: akhand@cse.kuet.ac.bd (M.A.H.A.); maskamal@gunma-u.ac.jp (M.A.S.K.)

Abstract: Electroencephalogram (EEG) has emerged as the most favorable source for recognizing brain disorders like epileptic seizure (ES) using deep learning (DL) methods. This study investigated the well-performed EEG-based ES detection method by decomposing EEG signals. Specifically, empirical mode decomposition (EMD) decomposes EEG signals into six intrinsic mode functions (IMFs). Three distinct features, namely, fluctuation index, variance, and ellipse area of the second order difference plot (SODP), were extracted from each of the IMFs. The feature values from all EEG channels were arranged in two composite feature forms: a 1D (i.e., unidimensional) form and a 2D image-like form. For ES recognition, the convolutional neural network (CNN), the most prominent DL model for 2D input, was considered for the 2D feature form, and a 1D version of CNN was employed for the 1D feature form. The experiment was conducted on a benchmark CHB-MIT dataset as well as a dataset prepared from the EEG signals of ES patients from Prince Hospital Khulna (PHK), Bangladesh. The 2D feature-based CNN model outperformed the other 1D feature-based models, showing an accuracy of 99.78% for CHB-MIT and 95.26% for PHK. Furthermore, the cross-dataset evaluations also showed favorable outcomes. Therefore, the proposed method with 2D composite feature form can be a promising ES detection method.

Keywords: epileptic seizure; electroencephalogram; empirical mode decomposition; 1D and 2D composite feature; convolutional neural network

1. Introduction

Epileptic seizure (ES) is a seizure when abnormal electrical activity arises in the brain due to the underlying condition of epilepsy. It is one of the most common neurological diseases worldwide [1] and is a major health concern. ES is usually diagnosed based on an individual's seizure patterns, medical history, and results of neurological exams and imaging tests. ES recognition is important because it tends to recur, and the potential consequences can be severe. As it is a brain disorder, detection through analyzing brain signals is the most favorable method, and electroencephalogram (EEG), an affordable, noninvasive method for measuring brain activity, is the most promising for ES detection.

Epilepsy has been a subject of extensive research in the computational intelligence domain over the last few decades ([2–6]) for automated ES detection and diagnosis. In recent years, machine learning (ML) and deep learning (DL) techniques have emerged as powerful tools for analyzing EEG for the diagnosis of neurological disorders such as autism and emotion [7,8]. EEG signals are the most promising brain signals for ES analysis and recognition, particularly for recognizing the abnormality of the brain due to its painless experiment and inexpensive nature [9]. EEG signals during seizure typically show abnormal rhythmic activity in the brain, known as spikes and sharp wave discharges.

EEG signals consist of some challenges regarding nonstationary and nonlinearity behavior, similarity in frequency and amplitude in seizure and non-seizure segments, and noise while recording [10].

Existing EEG-based ES studies have explored various techniques to process the EEG signals, extract relevant features, and use different ML/DL methods for ES detection. Some studies have focused on using raw EEG signals (e.g., [4–6,11–13]) and employed DL learning methods for feature extraction with embedded methods of individual DL methods. Furthermore, recent studies have explored various feature extraction methods to capture discriminative information from raw EEG signals [3,11,14–17]. Recent EEG-based ES studies have focused on signal transformation and feature extraction techniques to achieve better ES recognition using ML/DL methods. Some other studies have incorporated signal transformation techniques through decomposition and filtering [2,3,14].

The study aimed to investigate the well-performing ES detection method from EEG signals by considering these main steps: EEG signal decomposition, feature extraction, and the classification of ES using ML and DL methods with an appropriate representation of features. The empirical mode decomposition (EMD) scheme decomposes EEG signals iteratively into intrinsic mode functions (IMFs) (i.e., decomposed signals), where initial IMFs hold most of the information, and the information quality gradually degrades in the later IMFs. Three distinct features, namely, the fluctuation index, variance, and ellipse area of the second order difference plot (SODP), are extracted from the first six IMFs (i.e., 18 (=6 × 3) extracted features per EEG channel). The feature values from all EEG channels were arranged in two different composite feature forms: a 1D (i.e., unidimensional) form that places the channel features one after another in a sequence and a 2D image-like one that places the individual channel features in rows. A convolutional neural network (CNN), the most prominent DL model for 2D input, was considered for the 2D feature form, whereas a neural network (NN) and 1D version of the CNN were used for the 1D feature form for ES detection. The experimental evaluation and analysis were performed on the CHB-MIT benchmark EEG dataset and a dataset prepared from the EEG signals of ES patients from Prince Hospital Khulna (PHK), Bangladesh. Additionally, cross-dataset evaluations assessed the generalization and robustness of the models. Collectively, the significance of this study lies in the decomposition of EEG signals, feature extraction from IMFs, and reshaping them into two different forms, and the employment of appropriate ML/DL methods for ES classification. More specifically, the main contributions in the context of ES detection and the achievements of this study are:

1. The use of EMD for IMF extraction from the CHB-MIT and PHK datasets as well as the fluctuation index, variance, and ellipse area of SODP feature extraction from IMFs.
2. Representing the feature values in the 1D form and 2D image-like form and employing the appropriate CNN model for ES classification.
3. The model with 2D feature and CNN was identified as the most promising method for ES recognition, outperforming other approaches for the CHB-MIT and PHK datasets.
4. Cross-dataset evaluations and outcomes comprising other prominent studies revealed the proficiency of the proposed method.

The rest of the paper is structured as follows. Section 2 briefly reviews different ES studies using ML and DL methods. Section 3 describes the proposed method, which includes signal decomposition, feature extraction, feature reshaping the 1D/2D composite form, and classification using CNN. Section 4 includes the experimental studies, presents the experimental results and analysis, and compares the proposed method's outcomes with the state-of-art. Finally, Section 5 first summarizes the contribution and achievement of the present study, and then outlines several future research directions that have emerged from this work.

2. Related Works

According to the recent literature, various EEG-based ES detection methods use diverse techniques to process the EEG signal, extract features, and classify ES using different

ML/DL methods. An approach utilizing mutual information (MI), CNN, and learned factor graphs was investigated in [18] to detect seizures, where MI was used to find the correlation between channels for the CHB-MIT dataset. CNN (i.e., 2D CNN) was used to detect seizures from the CHB-MIT dataset without feature extraction [12]. Imaged EEG signals from CHB-MIT were also classified using CNN [13]. CNN was also used in [19] for seizure classification using plotted EEG images formed from EEG signals. 1D CNN, a variant of CNN, was used in [4] for ES detection by using the raw EEG of the Bonn dataset. In another study [5], a transfer learning concept was used in the ES detection process, where a pre-trained AlexNet CNN model was used on raw EEG data.

In [6], graph attention networks (GATs) were used for spatial feature extraction, and bi-directional long short-term memory (BiLSTM) was used for ES detection from the raw EEG signals of the CHB-MIT and TUH [20] datasets. The study in [21] investigated the effects of MinMaxScaler normalization and evaluated the outcomes obtained by employing recurrent neural network (RNN), LSTM, and BiLSTM models on the preprocessed CHB-MIT dataset. A method was proposed in [11] to deal with imbalanced seizure data in CHB-MIT: a generative adversarial network (GAN) was used for data enhancement, and a 1D CNN was trained on an augmented dataset containing original and generated EEG data. A deep LSTM network was proposed in [22] to learn the temporal dependencies using single-channel EEG data from the Bonn dataset. A lightweight 1D CNN-based ES detection model, LightSeizureNet, was introduced for the CHB-MIT dataset, featuring both patient-independent and patient-specific versions [23]. In [24], the CHB-MIT dataset was employed for ES detection, utilizing frequency bands to make brain networks in conjunction with CNN techniques.

Some studies have employed signal decomposition in ES recognition. EMD was used to decompose EEG signals to extract different types of entropy, fractal dimension, statistical, and exponential energy from the IMFs [2]. The study used a support vector machine (SVM) classifier to differentiate the seizure and non-seizure signals from the Bern-Barcelona EEG dataset. Cross-bi-spectrum EEG signal analysis and three linear and six nonlinear features were used [15], and SVM was used for ES classification from the Freiburg iEEG database [25]. SVM was also used to detect neonatal seizure in [16] with the time and time–frequency domain correlation features from the Helsinki University Hospital dataset. The Tunable Q-wavelet transform (TQWT)-based decomposition of CHB-MIT, combined with feature extraction (i.e., nonlinear, temporal, statistical) and classification using SVM and random forest (RF), were performed in ES classification in [3]. Genetic algorithm and particle swarm optimization were used to refine the parameters of hybrid SVM in [26] for the Bonn dataset. In [17], two time-domain feature extraction methods were utilized along with different classifiers like SVM, K nearest neighbors (KNN), logistic regression (LR), RF, decision tree (DT), naive Bayes (NB), etc., to detect seizure. The feedforward NN was used in [14] to classify seizures, where wavelet decomposition was used along with feature extraction. In [27], variational modal decomposition (VMD) was applied to decompose EEG signals from CHB-MIT, where features like differential entropy (DE) and high-frequency detection (HFD) were derived and used with an SVM classifier for ES detection. In another study [28], CNN was used to decompose EEG along with an ES detection model utilizing adversarial training and the attention-based CNN algorithm to achieve patient-independent diagnosis from raw EEG signals.

3. Epileptic Seizure Detection from EEG Using EMD and DL

Existing ML/DL-based methods are diverse in using raw EEG signal data with feature extraction from raw or decomposed EEG signals. In the case of decomposition, EMD can adapt to the local characteristics of a signal by dividing it into its intrinsic modes, which are the most significant and relevant parts of the signal. EMD-based decomposition was better than discrete wavelet transform (DWT) in [2]. Different features (e.g., entropy, fractal dimension, statistical, exponential energy) were extracted from the decomposed signals (i.e., IMFs), and classification using SVM from the features in the unimodal 1D form showed

remarkable performance in [5]. It is possible to use CNN, the most popular DL method, by reshaping the extracted features into a 2D form in the exploration for further improvements in detection performance.

The study intends to improve EEG-based ES recognition performance by developing an EMD and CNN model. The proposed framework of ES detection is illustrated in Figure 1, depicting the principal operation steps: preprocessing of the collected EEG signal, signal decomposition using EMD, feature extraction from decomposed signals (i.e., IMFs), composite feature formation through feature representation, and finally, classifying ES using ML and DL methods. Each signal channel produces three IMF samples through decomposition in the first stage. Next, three distinct features (i.e., fluctuation index, variance, and ellipse area of SODP) are extracted from each IMF. Therefore, there are eighteen extracted features for a single EEG channel. The composite feature is formed through feature representation, aggregating features for individual channels and then combining them for all channels. The classification of ES is finally performed from the composite feature with the appropriate ML/DL method. The following subsections describe the individual operational steps of the framework.
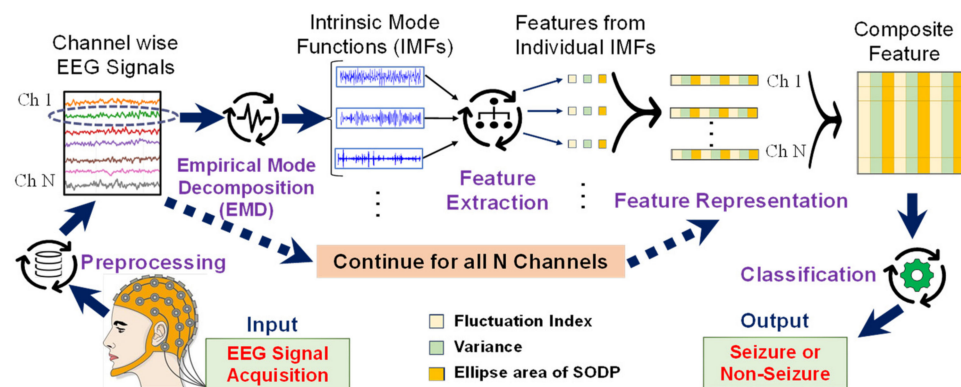
Figure 1. Proposed framework for epileptic seizure detection.

3.1. Data Collection and Preprocessing

The study used a publicly available CHB-MIT benchmark dataset and a locally collected EEG dataset from Prince Hospital Khulna (PHK), Bangladesh. Two datasets are described below.

3.1.1. CHB-MIT Dataset

The well-studied CHB-MIT dataset [29] was created by a team of investigators from the Children's Hospital Boston (CHB) and the Massachusetts Institute of Technology (MIT). In this dataset, the sampling rate of the signals is 256 samples per second. This study considered EEG signal data from 22 channels: FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8, P8-O2, FZ-CZ, CZ-PZ, P7-T7, T7-FT9, FT9-FT10, and FT10-T8. EEG data from 24 available cases were analyzed to prepare seizure and non-seizure samples. With a 10-s window size and 70% overlapping (i.e., 3-s step size), the EEG signal is divided into consecutive segments. Such segmentation ensures that no information is lost during the signal processing and allows for a more robust EEG signal analysis, particularly in cases where the activity of interest (e.g., seizures) may occur within smaller time frames. A total number of 3707 segments for seizure and 3707 segments for non-seizure were created. Therefore, 7414 collective samples were prepared for ML/DL for further steps.

3.1.2. Prince Hospital Khulna (PHK) Dataset

The EEG signal of the PHK dataset was recorded under the supervision of a specialist in neurology and epilepsy (i.e., an author of the study). Specifically, EEG signal data from 10 patients were taken into consideration in this study with approval from the hospital

authority (Ref.: PDK/2023/0125). Table 1 provides a concise summary of the patients in terms of age and gender; the patients' names, addresses, and other personal information have been kept anonymous for privacy purposes and the ethical clearance criteria set by the hospital. The Nicolet vEEG System manufactured by Natus Neuro, USA [30] was used to record the data; the model is NicOne, version 5.94. Different montage settings have been interpreted such as monopolar and bipolar (longitudinal and transverse). The recorded channels (in monopolar or referential montage) were FP1-A1, F3-A1, C3-A1, P3-A1, O1-A1, FP2-A2, F4-A2, C4-A2, P4-A2, O2-A2, F7-A1, T3-A1, T5-A1, F8-A2, T4-A2, T6-A2, FZ-A1, CZ-A1, PZ-A1, ECG-Bipolar, and Photic. Excluding 'ECG-Bipolar' and 'Photic', signals from 19 general channels were considered for processing for ES detection. The PHK EEG data are in .e(Nicolet) format, with 500 samples per second. EEGLAB toolbox (version 2021.1) of MATLAB was used to convert the Nicolet into CSV files and resample them to 256 Hz to make them compatible with CHB-MIT. EEG Viewer software (v5.71.3.2522) was used to visualize the signals and spikes of different channels. From 10 cases, 578 s of seizure signals and 578 s of non-seizure signals were collected. Notably, 380 segments (190 segments for seizure and 190 for non-seizure) were created using a 10-s window and 70% overlapping.

Table 1. Summary of cases of the PHK dataset.

Case No.	1	2	3	4	5	6	7	8	9	10
Age (Year)	28	9	13	6.5	18	4	3	7	17	1.5
Gender	F	M	F	M	M	M	F	F	M	M

3.2. EEG Signal Decomposition Using Empirical Mode

Signal decomposition enables a more in-depth analysis as simpler components can reveal underlying patterns. EMD decomposes EEG signals into intrinsic mode functions (IMFs), which represent different time-scale components of the signal. A detailed description of EMD is available in [2]. EMD iteratively extracts the IMFs from the signal until a residue is obtained that cannot be further decomposed; initial IMFs hold most of the information, and the information quality gradually degrades in the later IMFs. In this study, the first six IMFs were considered to carry sufficient information.

For a better understanding, the first six IMFs of the seizure and non-seizure segments in the PHK are shown in Figure 2. It can be observed from the figure that individual IMFs are also continuous signal-like time series. During seizures, EEG signals tend to exhibit sharp and irregular activities. The corresponding IMFs may contain high-frequency components with large amplitudes (see IMF4), representing the rapid and chaotic neuronal firing associated with seizures. In non-seizure EEG signals, the IMFs are generally smoother and more regular. The frequency components are mostly well-organized and follow predictable patterns, reflecting normal brain activity.
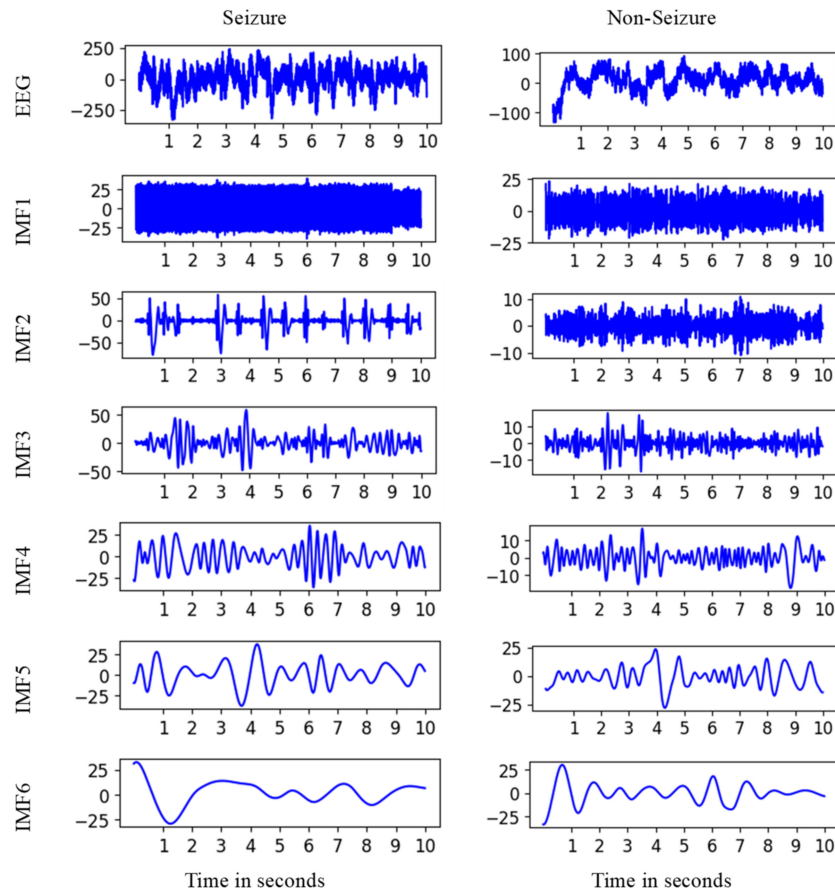
Figure 2. Six IMFs of seizure and non-seizure sample cases from the PHK dataset.

3.3. Feature Extraction from IMFs

In this study, features were extracted from six IMFs individually. From each IMF, three features were extracted. Short descriptions of the feature extraction methods are given below:

3.3.1. Fluctuation Index (FI)

The fluctuation index provides a measure of the relative level of variability in the data. FI can be used to compare different time series (here, an IMF) to identify data patterns and trends, which is obtained as:

$$\text{Fluctuatuion Index} = \frac{1}{X-1}\sum\nolimits_{x=0}^{X-2} |s_i(x+1) - s_i(x)|, \tag{1}$$

where X is the length of signal $s_i(x)$, here an IMF. A higher FI indicates that the data are more variable, while a lower value indicates that the data are more stable.

3.3.2. Variance

Variance is a measure of how far the individual data points in a set are from the mean or average of the data, which is obtained as:

$$\text{Variance} = \frac{\sum_{x=0}^{X-1} (s_i(x) - \mu)^2}{X-1}, \tag{2}$$

where μ is the mean of the signal $s_i(x)$.

3.3.3. Ellipse Area of Second Order Difference Plot (SODP)

Dispersion or variability of a time series dataset is measured by the ellipse area of SODP. The SODP is a graphical depiction of the second-order statistics of a time series, which can be used to spot trends and outliers in the data. The ellipse area of the SODP is considered an important feature for time series data because it provides a compact representation of the variability and spread of the data. The SODP is obtained as [31]:

$$d_1(x) = s_i(x + 1) - s_i(x), \tag{3}$$

$$d_2(x) = s_i(x + 2) - s_i(x - 1), \tag{4}$$

where $s_i(x)$ is the time series signal, here an IMF. To measure the SODP of the signal $s_i(x)$, $d_2(x)$ was plotted against $d_1(x)$. The visual analysis revealed that the initial IMFs of the seizure data exhibited a greater dispersion of data points in an elliptical pattern. The ellipse area of SODP is defined by:

$$\text{Area}_{\text{ellipse}} = \pi ab, \tag{5}$$

where a and b denote the semi-major and semi-minor axes, respectively. The semi-major and semi-minor axes with the parameters are given as:

$$p_1 = \sqrt{\frac{1}{X - 2}\sum_{x=0}^{X-2} d_1(x)^2}, \tag{6}$$

$$p_2 = \sqrt{\frac{1}{X - 2}\sum_{x=0}^{X-2} d_2(x)^2}, \tag{7}$$

$$p_3 = \frac{1}{X - 2}\sum_{x=0}^{X-4} d_1(x) \times d_2(x), \tag{8}$$

$$\delta = \sqrt{p_1{}^2 + p_2{}^2 - 4\left(p_1{}^2 \times p_2{}^2 - p_3{}^2\right)}, \tag{9}$$

$$a = 1.7321\sqrt{p_1{}^2 + p_2{}^2 + \delta}, \tag{10}$$

$$b = 1.7321\sqrt{p_1{}^2 + p_2{}^2 - \delta} \tag{11}$$

where X is the length of $s_i(x)$.

3.4. Feature Representation

Composite feature formation representing the extracted ones is necessary to input them perfectly in the ML/DL model for ES classification, which is the crucial issue of this study. Each channel has six IMFs, and each IMF is further decomposed into three features, i.e., making 18 (=3 × 6) features per channel. It is essential to carefully consider the characteristics of the features and select the most appropriate representation technique to ensure the optimal ML/DL model performance. This study organized the features in two formats, as illustrated in Figure 3: 1D or unidimensional and 2D forms. These two different forms of organizing the feature values offer different perspectives for analyzing EEG data. In the 1D form represented in Figure 3a, the 18 feature values of each EEG channel are placed one after another in a single line. This arrangement creates a linear sequence where the feature values of the channels are concatenated. For instance, the feature values of 22 EEG channels of CHB-MIT were arranged in a continuous sequence of 396 (=18 × 22) values. For the PHK dataset, it was a continuous sequence of 342 (=18 × 19) values for 19 EEG channels.
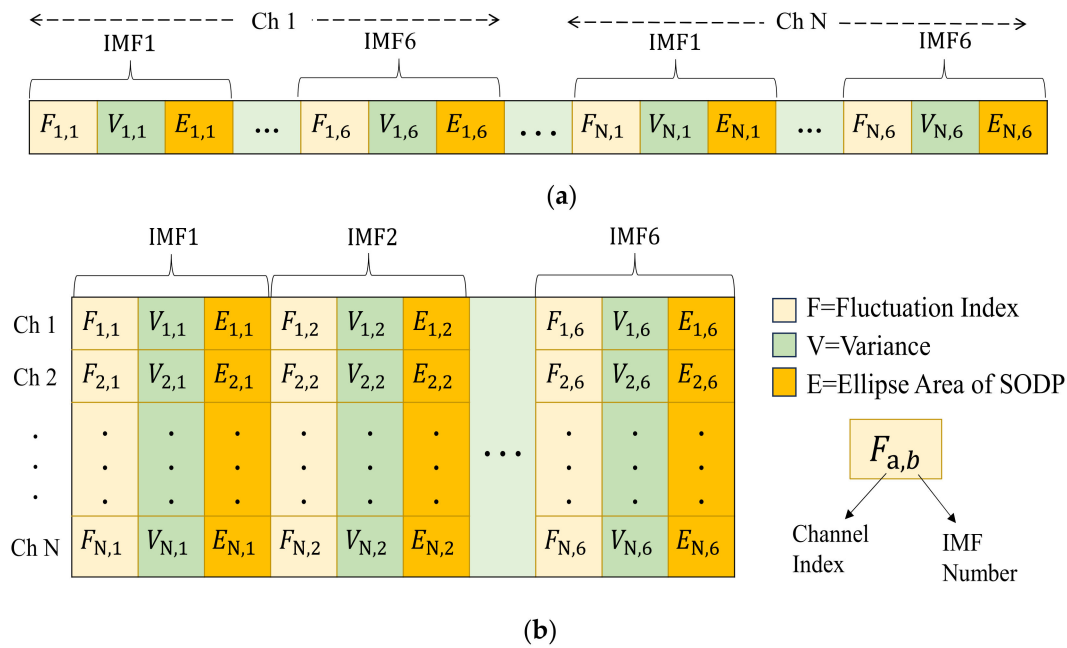
Figure 3. Two different feature orientations of 18 features (= 6 IMFs × 3 features) for each individual channel (Ch). Total channels (N) were 22 and 19 for the CHB-MIT and PHK datasets, respectively. (**a**) All of the features are placed in a line placing individual channel features one after another in a line. (**b**) The feature values of individual channels are placed in rows in the 2D image-like form.

The feature values of individual EEG channels were placed in rows to represent features in 2D form. The resulting structure resembled an image-like form with values in a 2D matrix, where each row corresponds to a channel, and the columns represent different features. A particular row in the matrix holds 18 feature values (i.e., fluctuation index (F), variance (V), and ellipse area of SODP (E) features) for six IMFs of an EEG channel signal value. At a glance, for the CHB-MIT dataset, the features of 22 channels were placed in 22 rows with dimensions of 22 × 18. In the matrix depicted in Figure 3b F1,1 in [1,1] position is the fluctuation index feature value of Channel 1's (i.e., Ch 1) IMF1. The next two values, V1,1 and E1,1, are the variance and ellipse area of the SODP features for the same IMF. The last value in the first row E1,6 in position [1,18] was the ellipse area of the SODP feature for the IMF6 of Channel 1. As N = 22 for the CHB-MIT dataset, the remaining 21 rows hold feature values for Ch 2 to Ch 22. Similarly, N = 19 for the PHK dataset; therefore, the feature values of the channels were placed in 19 individual rows in a 19 × 18 size in 2D form.

3.5. Seizure Classification

Appropriate seizure detection from composite features in 1D or 2D forms using suitable classifiers is essential. The following subsections describe the architectures of the three classifiers considered in this study.

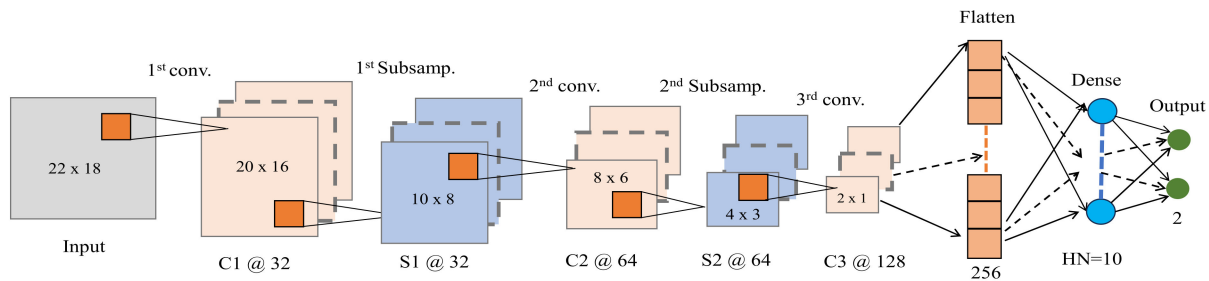### 3.5.1. 2D Feature Classification with CNN

CNN is a well-studied DL model to classify image or image-like 2D representations of features ([7,8]) due to its unique ability to process and analyze 2D structured data. Convolution operation with a 2D kernel makes CNN (hereafter called 2D CNN) particularly effective in recognizing patterns, shapes, and structures in images or 2D input. Additionally, pooling is employed to subsample the extracted features, reducing their size while retaining their essential properties. After several successful convolution and pooling operations, classification is finally conducted through the dense layer.

For the 22×18 size composite feature (depicted in Figure 3b) for CHB-MIT, the CNN architecture with three convolution layers (C1, C2, and C3), two pooling layers (S1 and S2), and dense layer with HN = 10 is

$$I_{22 \times 18} \rightarrow \{32K1_{3 \times 3}\,C1_{20x16} - S_{2 \times 2}32S1_{10 \times 8}\} \rightarrow \{64K2_{3 \times 3}C2_{8 \times 6} - S_{2 \times 2}64S2_{4 \times 3}\}$$

$$\rightarrow \{128K3_{3 \times 3}C3_{2 \times 1}\} \rightarrow \{Wo_{256 \times 10}\} \rightarrow \{Wo_{10 \times 2}\} \rightarrow O_2$$

Figure 4 illustrates the model architecture. In the model, the size of the kernels K1, K2, and K3 is 3×3 with stride 1, but the number of filters increases progressively from 32, 64, and 128. After the first convolutional layer, batch normalization is applied to normalize the activations. The max pooling areas for S1 and S2 are 2×2 with stride 2. The final sub-sampled feature maps are placed linearly into a vector, preparing it for the fully connected dense layer to classify seizure and non-seizure with two output nodes O2. Activation functions in dense layers are ReLU and SoftMax in the hidden and output layers, respectively. For HN = 50, the sizes of dense layer weights are updated as $\{Wo_{256 \times 50}\} \rightarrow \{Wo_{50 \times 2}\}$. For the PHK dataset with 19 channels, the input was $I_{19 \times 18}$, and other shapes were updated accordingly.

Figure 4. CNN structure to classify seizure from a 22×18 sized composite feature for the CHB-MIT dataset.

3.5.2. 1D Feature Classification with 1D CNN

A 1D form of CNN was used for the 1D shape composite feature in Figure 3a. The 1D CNN is a variant of 2D CNN with convolutional/pooling features and filters in 1D form. For a total of 396, CHB-MIT features with HN = 10 is

$$I_{396 \times 1} \rightarrow \{32K1_{3 \times 1}\,C1_{394 \times 1} - S_{2x1}32S1_{197 \times 1}\} \rightarrow \{64K2_{3 \times 1}C2_{195 \times 1} - S_{2 \times 1}64S2_{97 \times 1}\}$$

$$\rightarrow \{128K3_{3 \times 1}C3_{95 \times 1} - S_{2 \times 1}128S3_{47 \times 1}\} \rightarrow \{256K4_{3 \times 1}C4_{45 \times 1} - S_{2 \times 1}256S4_{22 \times 1}\}$$

$$\rightarrow \{Wo_{5632 \times 10}\} \rightarrow \{Wo_{10 \times 2}\} \rightarrow O_2$$

It has four convolutional layers with increasing filter sizes: 32, 64, 128, and 256. For the PHK dataset with 342 features, the input was $I_{342 \times 1}$ and other shapes were updated accordingly.

4. Experimental Studies

The proposed seizure classification methods were tested on the CHB-MIT and PHK datasets to evaluate their performance. Aside from using 2D and 1D CNNs, an NN model with a hidden layer with HN 10 and 50 was considered to classify 1D features to better understand their comparative performances. In addition, cross-dataset evaluation was also conducted.

4.1. Experimental Setup

In this study, $MATLAB^{®}$ R2021a was used to preprocess and resample the PHK dataset through the device with the following configuration: CPU: 11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40 GHz, RAM: 8.00 GB, 64-bit Windows operating system. The Python programming language was used to conduct the experiments. Decomposition and

feature extraction, with model training, were powered by a P100 GPU of Kaggle's platform. Fivefold cross-validation (CV) was used to split the training and test sets of the available 7414 segments of the CHB-MIT dataset and 380 segments of the PHK dataset. In a 5-fold CV, the available samples are divided into five portions, where a portion is considered a test set by tern and the remaining portions are used to train the model. Adam optimization was used with a learning rate of 0.001. The classifier models used HN values in dense layers for 10 and 50.

4.2. Evaluation of CHB-MIT Dataset

Figure 5 shows the training loss curves for a sample fold case (i.e., Fold 1) for three different ML and DL models for HN = 10 and HN = 50. It can be observed from the figure that CNN (both 1D and 2D) had faster convergence than NN, and the phenomena were clearly visible up to 20 epochs. However, all six models converged after 50 epochs. Figure 6 shows the test set accuracy curves for the six models, whose loss curves are presented in Figure 5 (i.e., for Fold 1). Significant differences among the models were found in the test set accuracies.
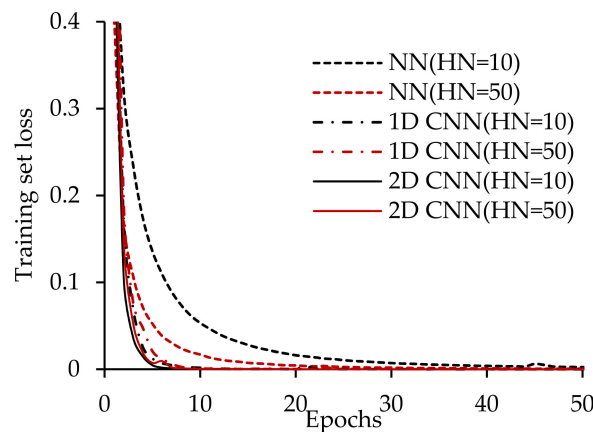
Figure 5. Training set loss vs. epochs for the CHB-MIT dataset.
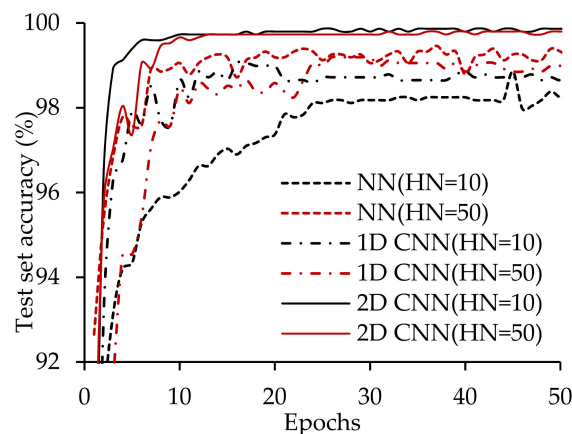
Figure 6. Test set accuracy vs. epochs for the CHB-MIT dataset.

Faster test set accuracy improvement with epochs was shown for 2D CNN, and the worst improvement was observed for NN. Based on the overall test accuracy, 2D CNN was the best, and NN was the worst.

Table 2 compares the test set accuracies for five individual folds and an average of the folds among the six models. Based on the average accuracy, 2D CNN (HN = 50) showed the highest accuracy of 99.78%. This was closely followed by the 2D CNN model with HN = 10, which achieved an accuracy of 99.71%. The 1D CNN model also performed well, reaching 99.51% and 99.58% accuracies for HN = 10 and HN = 50, respectively. These

results emphasize the suitability of the 1D CNN architecture. The traditional NN models were found to be inferior to the 1D CNN and 2D CNN models, as NN (HN = 50) and NN (HN = 10) achieved accuracies of 99.09% and 99.39%, respectively. These results highlight the effectiveness of the 2D CNN architecture in capturing distinguishing patterns within the data and achieving high accuracy levels.

Table 2. Test set accuracy of the 5-fold cross validation of the CHB-MIT dataset.

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Avg. of 5-Fold CV
NN (HN = 10)	99.06	99.39	99.19	98.85	98.99	99.09
NN (HN = 50)	99.53	99.46	99.39	99.33	99.26	99.39
1D CNN (HN = 10)	99.53	99.60	99.39	99.66	99.39	99.51
1D CNN (HN = 50)	99.6	99.66	99.46	99.73	99.46	99.58
2D CNN (HN = 10)	99.93	99.66	99.6	99.73	99.66	99.71
2D CNN (HN = 50)	99.93	99.80	99.73	99.73	99.73	99.78

4.3. Evaluation of PHK Dataset

The PHK dataset was evaluated in the same manner as the CHB-MIT dataset presented in the previous section. Figure 7 shows the training loss curves for a sample fold case (i.e., Fold 3 because it showed the variations of models properly) for three different ML and DL models for HN = 10 and HN = 50. It was observed from the figure that the CNN (both 1D and 2D) had faster convergence than the NN, and the phenomena were clearly visible up to 100 epochs. However, all six methods converged after 200 epochs. Figure 8 shows the test set accuracy curves for the six models. It is remarkable, as seen in Figure 6, that the accuracy curves of the PHK dataset were not smooth like those of the CHB-MIT dataset. However, after 150 epochs, all models were near convergence. Significant differences among the models with many fluctuations over epochs were found in the test set accuracies. The smaller number of samples in the PHK dataset was the main reason for such phenomena. At a glance, 2D CNN was the best and NN was the worst, based on the overall test accuracy.
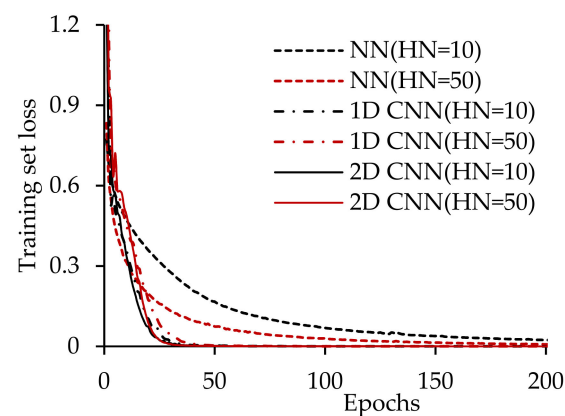
Figure 7. Training set loss vs. epochs for the PHK dataset.
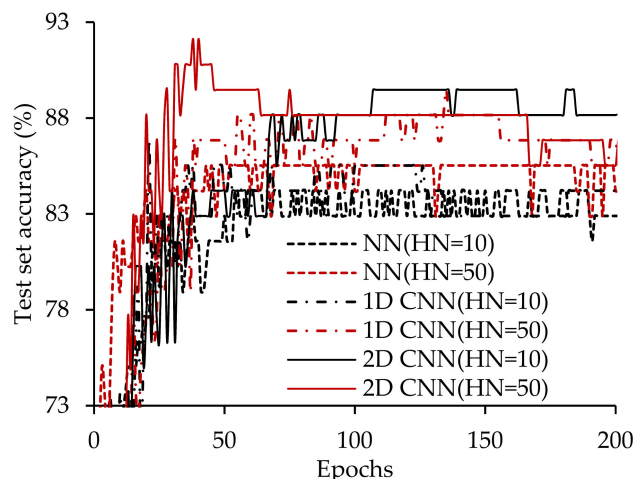
Figure 8. Test set accuracy vs. epochs for the PHK dataset.

Table 3 compares the test set accuracies for five individual folds and an average of the folds among the six models. Based on average accuracy, NN (HN = 10) showed a 90.79% accuracy, and NN (HN = 50) achieved a slightly better accuracy of 91.05%. Moving to the CNN architecture, 1D CNN (HN = 10) demonstrated improved accuracy, reaching 91.84%. Subsequently, 1D CNN (HN = 50) further enhanced the accuracy to 92.63%. The performance improvement was notable when switching to the 2D CNN models. 2D CNN (HN = 10) exhibited a noteworthy accuracy of 93.94%. Exceeding the performance of all other models, 2D CNN (HN = 50) showcased the best accuracy of 95.26%. Notably, the 2D CNN models consistently outperformed their counterparts, suggesting that their ability to capture patterns led to a superior classification performance.

Table 3. Test set accuracy of the 5-fold cross validation of the PHK dataset.

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Avg. of 5-Fold CV
NN (HN = 10)	92.11	88.16	84.21	93.42	96.05	90.79
NN (HN = 50)	90.79	90.79	85.53	92.11	96.05	91.05
1D CNN (HN = 10)	92.11	92.11	86.84	92.11	96.05	91.84
1D CNN (HN = 50)	93.42	93.42	89.47	90.79	96.05	92.63
2D CNN (HN = 10)	94.74	93.42	89.47	96.05	96.05	93.94
2D CNN (HN = 50)	96.05	93.42	92.11	97.37	97.37	95.26

4.4. Evaluation of Cross-Dataset

Cross-dataset evaluation is vital as it ensures that ML models can generalize their knowledge, handle variations, avoid biases, and perform reliably in real-world scenarios. In a remarkable effort, this study conducted a cross-dataset evaluation for better understanding, especially for the locally obtained PHK dataset, which differs from the CHB-MIT datasets obtained under different environmental conditions. Furthermore, there are variations in the data collection methodologies and instrumentations. From both datasets, 12 common channels were taken, which were FP1, F3, C3, P3, FP2, F4, C4, P4, F7, F8, FZ, and CZ. Although only 12 common channels were taken for the cross-dataset evaluation, there were some mismatches as the reference channels of both datasets were different. The total number of features for 12 channels was 216 (=12 × 18). The feature vectors were reshaped as 12 × 18 for the 2D CNN model. Based on the input size of 12 × 18, the 2D CNN architecture was changed to two convolutional layers and one max pooling layer. First, the models were trained with all 7414 segments of the CHB-MIT dataset and tested with all 380 segments of the PHK dataset. Second, the models were trained with the PHK and

tested with the CHB-MIT. The training set loss curves were almost similar to the individual datasets presented in the previous sections; therefore, they are not discussed here for the sake of brevity.

The main concern of cross-dataset evaluation is the generalization ability. Table 4 shows the overall accuracy of the cross-dataset evaluation for different ML and DL models. In the test performance of the PHK dataset, the highest accuracy was 74.21%, achieved by 1D CNN (HN = 50). On the other hand, the best test performance of the CHB-MIT dataset was 64.12% by 2D CNN (HN = 10), which was lower than the PHK dataset by any model. The CHB-MIT dataset comprises 7414 samples of diverse features, providing a rich training dataset that allowed the model to learn from a wide range of data. The accuracy of a model trained by CHB-MIT indicates its ability to generalize well due to a diverse range of features. In contrast, the PHK dataset holds only 380 samples, and potentially lacks in limiting the model's ability to generalize effectively. As a result, the test performance of the CHB-MIT dataset while the model trained with PHK showed a relatively low accuracy for any ML/DL model. Another issue with the low accuracy on the cross-dataset evaluation was the lower number of channels, as 12 common channels were considered in this study. However, the results presented in Table 4 reveal that a cross-dataset is effective for real-life local cases when samples are limited.

Table 4. Cross performance of the CHB-MIT dataset and PHK dataset.

Model	Test Performance of PHK (Training with CHB-MIT)	Test Performance of CHB-MIT (Training with PHK)
NN (HN = 10)	66.58	61.15
NN (HN = 50)	66.84	61.67
1D CNN (HN = 10)	71.32	62.17
1D CNN (HN = 50)	74.21	60.36
2D CNN (HN = 10)	72.32	64.12
2D CNN (HN = 50)	69.74	63.84

4.5. Performance Comparison with Existing Studies

Several recent studies have conducted seizure detection with different ML/DL methods with the CHB-MIT dataset. The outcomes of these studies are summarized, and performance comparisons on the test set with the proposed method are demonstrated in Table 5. The results of the existing studies are the reported results in the corresponding articles. Some studies [12,13] used raw EEG signals without any feature extraction technique. Kaziha and Bonny [12] segmented the dataset by 100 s and achieved an accuracy of 96.70% using 2D CNN. Gómez et al. [13] also used 2D CNN on raw EEG signal data and achieved a 99.30% accuracy following the leave-one-patient-out evaluation strategy. Deepa and Ramesh [21] considered the Minmax scaled CHB-MIT dataset and achieved a 99.55% accuracy using BiLSTM. To exploit the spatial and temporal relationships of 16 channels of CHB-MIT, He et al. [6] used GAT and BiLSTM for seizure classification and reported an accuracy of 98.52% in 5-fold CV. Recently, Qiu et al. [23] achieved a 97.09% accuracy in 10-fold CV using 1D CNN.

Some studies have used signal decomposition techniques. Segmenting the dataset by 2 s, Pattnaik et al. [3] decomposed the EEG signal using TQWT, extracted different features, and showed a 93% accuracy in 10-fold CV by RF. Segmenting the dataset by 1 s, Dang et al. [24] decomposed the EEG signal into frequency bands and showed a 99.56% accuracy in 10-fold CV by 2D CNN. Compared to all of them, our method achieved the best accuracy of 99.78% for CHB-MIT in a 5-fold CV. The proposed method's outperformance revealed the signal decomposition's effectiveness using EMD, feature extraction, and classification using 2D CNN from the 2D representation of the feature values.

Table 5. Comparison of the proposed method with prominent existing studies on epileptic seizure detection using the CHB-MIT dataset.

Author [Ref.], Year	Segment Time (Overlap%)	Train–Test Split	Decomposition + Feature Extraction	Classification Using ML/DL	Achieved Accuracy (%)
Kaziha and Bonny [12], 2020	100 s (No overlap)	70/30	N/A (used raw signal)	2D CNN	96.70
Gómez et al. [13], 2020	4 s (No overlap)	Leave-one-patient-out	N/A (used raw signal)	2D CNN	99.30
Dang et al. [24], 2021	1 s (50% overlap)	10-fold CV	Frequency bands + N/A	2D CNN	99.56
Pattnaik et al. [3], 2022	2 s (No overlap)	10-fold CV	TQWT + nonlinear, temporal, statistical feature	RF	93.00
He et al. [6], 2022	1 s (50% overlap)	5-fold CV	N/A + GAT	BiLSTM	98.52
Deepa and Ramesh [21], 2022	No segmentation	80/20	N/A + Minmax scaled	BiLSTM	99.55
Qiu et al. [23], 2023	2 s (50% overlap)	10-fold CV	N/A + 1D CNN	1D CNN	97.09
The proposed method	10 s (70% overlap)	5-fold CV	EMD + fluctuation index, variance, ellipse area of SODP	2D CNN	99.78

5. Conclusions

This study investigated a novel strategy of epileptic seizure (ES) detection from EEG signals through signal decomposition, feature extraction, and classification using DL methods with appropriate feature representation. The proposed method effectively uses EMD to decompose EEG signals into IMFs and extract pertinent features from six IMFs. The extracted feature representation in the 2D and 1D composite forms is the main technical novelty of the study. The different feature formations enhance the scope to employ various ML/DL models, and the study considered generic CNN (called 2D CNN) for the 2D feature form and 1D version of CNN and NN for the 1D feature form. Considering CHB-MIT (the popular benchmark dataset on ES) and the self-prepared dataset by processing EEG signal data from a local hospital, computational evaluations of the proposed method were conducted on intra-dataset and cross-dataset classifications. The 2D CNN with 2D feature form outperformed its counterparts (i.e., 1D CNN and NN) for both datasets in the intra-dataset classification. The proposed method was identified as the best-suited ES detection method compared to prominent, recent existing methods. On the other hand, achieving robust performance in cross-dataset evaluation remains a significant challenge in enhancing the generalization capabilities of real-life ES detection. Furthermore, the absence of channel reduction necessitates additional hardware setup, which presents a practical constraint.

In future work, the dimensions or structures of the decomposed features can be enhanced, and advanced DL models (e.g., 3D CNN) can be employed to investigate the possibility of further improvement with cross-dataset validations. Future studies can also explore ensemble techniques (e.g., bagging or boosting) to leverage the diversity of multiple ML/DL models using the base ML/DL models trained with individual feature types. Channel selection methods can be explored as a means to overcome hardware limitations without compromising the detection accuracy.

Author Contributions: M.A.H.A., S.D. and S.A.M. developed the concept. S.D. and S.A.M. conducted the experiments. M.A.H.A., S.D. and S.A.M. analyzed the results and prepared figures and tables. M.A.H.A. and S.D. prepared the initial manuscript draft, taking suggestions from M.A.S.K. and A.S. All authors contributed to editing and reviewing the manuscript. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Approval from Prince Hospital Khulna, Bangladesh (Ref.: PDK/2023/0125, Date: 25 September 2023).

Informed Consent Statement: Patients' diagnosis data (i.e., EEG signal) is used for computational model development only without exposing identity following the approval of the hospital authority.

Data Availability Statement: The well-known CHB-MIT EEG benchmark dataset and PHK local dataset were used in this study. The CHB-MIT dataset is freely available (https://physionet.org/content/chbmit/1.0.0/, accessed on 30 April 2024). PHK dataset is available from the corresponding author upon request. Source code and processed PHK dataset are available in the GitHub repository and the link is https://github.com/Shupta-das/Epileptic_seizure_detection_CNN, accessed on 24 April 2024.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Health Organisation. Epilepsy. 2022. Available online: https://www.who.int/news-room/fact-sheets/detail/epilepsy (accessed on 5 February 2023).
2. OK, F.; Rajesh, R. Empirical Mode Decomposition of EEG Signals for the Effectual Classification of Seizures. In *Advances in Neural Signal Processing*; IntechOpen: London, UK, 2020. [CrossRef]
3. Pattnaik, S.; Rout, N.; Sabut, S. Machine learning approach for epileptic seizure detection using the tunable-Q wavelet transform based time–frequency features. *Int. J. Inf. Technol.* **2022**, *14*, 3495–3505. [CrossRef]
4. Sameer, M.; Gupta, B. CNN based framework for detection of epileptic seizures. *Multimed. Tools Appl.* **2022**, *81*, 17057–17070. [CrossRef]
5. Nogay, H.S.; Adeli, H. Detection of Epileptic Seizure Using Pretrained Deep Convolutional Neural Network and Transfer Learning. *Eur. Neurol.* **2020**, *83*, 602–614. [CrossRef] [PubMed]
6. He, J.; Cui, J.; Zhang, G.; Xue, M.; Chu, D.; Zhao, Y. Spatial–temporal seizure detection with graph attention network and bi-directional LSTM architecture. *Biomed. Signal Process. Control* **2022**, *78*, 103908. [CrossRef]
7. Peya, Z.J.; Akhand, M.A.H.; Srabonee, J.F.; Siddique, N. EEG Based Autism Detection Using CNN Through Correlation Based Transformation of Channels' Data. In Proceedings of the 2020 IEEE Region 10 Symposium (TENSYMP), Dhaka, Bangladesh, 5–7 June 2020; pp. 1278–1281. [CrossRef]
8. Akhand, M.A.H.; Maria, M.A.; Kamal, M.A.S.; Murase, K. Improved EEG-based emotion recognition through information enhancement in connectivity feature map. *Sci. Rep.* **2023**, *13*, 13804. [CrossRef] [PubMed]
9. Hassan, K.M.; Islam, M.R.; Nguyen, T.T.; Molla, M.K.I. Epileptic seizure detection in EEG using mutual information-based best individual feature selection. *Expert Syst. Appl.* **2022**, *193*, 116414. [CrossRef]
10. Aayesha; Qureshi, M.B.; Afzaal, M.; Qureshi, M.S.; Gwak, J. Fuzzy-Based Automatic Epileptic Seizure Detection Framework. *Comput. Mater. Contin.* **2022**, *70*, 5601–5630. [CrossRef]
11. Gao, B.; Zhou, J.; Yang, Y.; Chi, J.; Yuan, Q. Generative adversarial network and convolutional neural network-based EEG imbalanced classification model for seizure detection. *Biocybern. Biomed. Eng.* **2022**, *42*, 1–15. [CrossRef]
12. Kaziha, O.; Bonny, T. A convolutional neural network for seizure detection. In Proceedings of the 2020 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, United Arab Emirates, 4 February–9 April 2020. [CrossRef]
13. Gómez, C.; Arbeláez, P.; Navarrete, M.; Alvarado-Rojas, C.; Le Van Quyen, M.; Valderrama, M. Automatic seizure detection based on imaged-EEG signals through fully convolutional networks. *Sci. Rep.* **2020**, *10*, 21833. [CrossRef]
14. Hassan, K.M.; Islam, M.R.; Tanaka, T.; Molla, M.K.I. Epileptic Seizure Detection from EEG Signals Using Multiband Features with Feedforward Neural Network. In Proceedings of the 2019 International Conference on Cyberworlds (CW), Kyoto, Japan, 2–4 October 2019; pp. 231–238. [CrossRef]
15. Mahmoodian, N.; Boese, A.; Friebe, M.; Haddadnia, J. Epileptic seizure detection using cross-bispectrum of electroencephalogram signal. *Seizure* **2019**, *66*, 4–11. [CrossRef]
16. Tapani, K.T.; Vanhatalo, S.; Stevenson, N.J. Time-Varying EEG Correlations Improve Automated Neonatal Seizure Detection. *Int. J. Neural Syst.* **2019**, *29*, 1850030. [CrossRef]
17. Nandini, D.; Yadav, J.; Rani, A.; Singh, V. Improved Patient-Independent Seizure Detection System Using Novel Feature Extraction Techniques. In *Advances in Intelligent Systems and Computing*; Springer: Singapore, 2022; pp. 879–888.
18. Salafian, B.; Ben-Knaan, E.F.; Shlezinger, N.; de Ribaupierre, S.; Farsad, N. CNN-Aided Factor Graphs with Estimated Mutual Information Features for Seizure Detection. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 23–27 May 2022; pp. 8677–8681. [CrossRef]
19. Emami, A.; Kunii, N.; Matsuo, T.; Shinozaki, T.; Kawai, K.; Takahashi, H. Seizure detection by convolutional neural network-based analysis of scalp electroencephalography plot images. *NeuroImage Clin.* **2019**, *22*, 101684. [CrossRef]
20. Obeid, I.; Picone, J. The Temple University Hospital EEG Data Corpus. *Front. Neurosci.* **2016**, *10*, 196. [CrossRef]
21. Deepa, B.; Ramesh, K. Epileptic seizure detection using deep learning through min max scaler normalization. *Int. J. Health Sci.* **2022**, *6*, 10981–10996. [CrossRef]
22. Hussein, R.; Palangi, H.; Ward, R.K.; Wang, Z.J. Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals. *Clin. Neurophysiol.* **2019**, *130*, 25–37. [CrossRef]
23. Qiu, S.; Wang, W.; Jiao, H. LightSeizureNet: A Lightweight Deep Learning Model for Real-Time Epileptic Seizure Detection. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 1845–1856. [CrossRef]

24.	Dang, W.; Lv, D.; Rui, L.; Liu, Z.; Chen, G.; Gao, Z. Studying Multi-Frequency Multilayer Brain Network via Deep Learning for EEG-Based Epilepsy Detection. *IEEE Sens. J.* **2021**, *21*, 27651–27658. [CrossRef]

25.	Ihle, M.; Feldwisch-Drentrup, H.; Teixeira, C.A.; Witon, A.; Schelter, B.; Timmer, J.; Schulze-Bonhage, A. EPILEPSIAE—A European epilepsy database. *Comput. Methods Programs Biomed.* **2012**, *106*, 127–138. [CrossRef]

26.	Subasi, A.; Kevric, J.; Canbaz, M.A. Epileptic seizure detection using hybrid machine learning methods. *Neural Comput. Appl.* **2019**, *31*, 317–325. [CrossRef]

27.	Sun, Y.; Chen, X. Epileptic EEG Signal Detection Using Variational Modal Decomposition and Improved Grey Wolf Algorithm. *Sensors* **2023**, *23*, 8078. [CrossRef]

28.	Zhang, X.; Yao, L.; Dong, M.; Liu, Z.; Zhang, Y.; Li, Y. Adversarial Representation Learning for Robust Patient-Independent Epileptic Seizure Detection. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2852–2859. [CrossRef] [PubMed]

29.	Shoeb, A. Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2009.

30.	NicoletOne^{TM} EEG System. Natus. Available online: https://natus.com/products-services/nicoletone-eeg-system (accessed on 22 June 2023).

31.	Pachori, R.B.; Patidar, S. Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions. *Comput. Methods Programs Biomed.* **2014**, *113*, 494–502. [CrossRef] [PubMed]